



Kernel estimation of extreme regression risk measures

Jonathan El Methni, Laurent Gardes, Stéphane Girard

► To cite this version:

Jonathan El Methni, Laurent Gardes, Stéphane Girard. Kernel estimation of extreme regression risk measures. *Electronic Journal of Statistics* , 2018, 12 (1), pp.359–398. 10.1214/18-EJS1392 . hal-01393519v3

HAL Id: hal-01393519

<https://inria.hal.science/hal-01393519v3>

Submitted on 14 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Kernel estimation of extreme regression risk measures

Jonathan El Methni⁽¹⁾, Laurent Gardes⁽²⁾ and Stéphane Girard^(3, *)

⁽¹⁾ *Université Paris Descartes, Sorbonne Paris Cité, MAP5, UMR CNRS 8145, France*

⁽²⁾ *Université de Strasbourg, CNRS, IRMA UMR 7501, France*

⁽³⁾ *Team Mistis, Inria Grenoble Rhône-Alpes & LJK, France*

Abstract

The Regression Conditional Tail Moment (RCTM) is the risk measure defined as the moment of order $b \geq 0$ of a loss distribution above the upper α -quantile where $\alpha \in (0, 1)$ and when a covariate information is available. The purpose of this work is first to establish the asymptotic properties of the RCTM in case of extreme losses, *i.e* when $\alpha \rightarrow 0$ is no longer fixed, under general extreme-value conditions on their distribution tail. In particular, no assumption is made on the sign of the associated extreme-value index. Second, the asymptotic normality of a kernel estimator of the RCTM is established, which allows to derive similar results for estimators of related risk measures such as the Regression Conditional Tail Expectation/Variance/Skewness. When the distribution tail is upper bounded, an application to frontier estimation is also proposed. The results are illustrated both on simulated data and on a real dataset in the field of nuclear reactors reliability.

Keywords: Conditional tail moment, Kernel estimator, Asymptotic normality, Risk measures, Extreme-value index, Extreme-value analysis.

AMS 2000 subject classification: 62G32, 62G30, 62E20.

1 Introduction

A recurrent problem in actuarial science, econometrics or statistical finance is to quantify the risk associated with a non-negative loss variable Y . A large variability of the random variable Y implies a high capital reserve for portfolios or a high price of the insurance risk. Quantiles are the basic tools in risk management and the main quantile-based risk measure in financial institutions is the Value at Risk with a confidence level $1 - \alpha$. It is defined as the α th quantile of the survival distribution of Y , see [35] for a review. When a covariate information X is recorded simultaneously with Y , the Value at Risk becomes a conditional quantile and is referred to as the Regression Value at Risk, denoted by $\text{RVaR}(\alpha|X)$ to emphasize the dependence on the covariate. The estimation of extreme Regression Value at Risk, *i.e.* $\text{RVaR}(\alpha|X)$ for small probabilities α , has many important applications, for instance in ecology [45], climatology [20], biostatistics [38], econometrics [9], finance [48], and insurance [4]. Recently, [11] extended the classical asymptotic theory on conditional

*Corresponding author, Stephane.Girard@inria.fr

quantiles [5, 44, 46, 47] further into the tails of the distribution by considering orders $\alpha = \alpha_n \rightarrow 0$ as the sample size n tends to infinity. The results are based on extreme-value theory [28], they hold true whatever the nature of the distribution tail. At the same time, an alternative regression risk measure, the Regression Conditional Tail Moment (RCTM) [16] was proposed to overcome the limitations of RVaR which prevent it from being a coherent risk measure [3]. The introduction of the RCTM permitted to adapt some risk measures to the regression setting, among them: the Conditional Tail Expectation (CTE) [3], also known as Tail-Value-at-Risk or Expected Shortfall, the Conditional Tail Variance (CTV) [49], the Conditional Tail Skewness (CTS) [32], etc. The authors also investigated the estimation of the RCTM for extreme levels within the context of heavy-tailed distributions.

The goal of this work is to fill in the gap between the two previous lines of work. Here, the asymptotic properties of the RCTM are established for extreme levels, and for all kinds of distribution tails. A nonparametric estimator is also introduced and its asymptotic distribution is derived in this context *i.e.* for $\alpha = \alpha_n \rightarrow 0$ as the sample size $n \rightarrow \infty$ and for an arbitrary distribution tail. As a first application, we obtain the asymptotic properties of the associated estimators: regression CTE, CTV and CTS. The second application takes place in the context of frontier estimation. Indeed, when the upper tail of the distribution of Y given $X = x$ is bounded, the right endpoint $y^*(x)$ is finite and is often referred to as a frontier. The estimation of $x \mapsto y^*(x)$ has received a lot of attention and various methods have been proposed based either on extreme-value estimators [22, 25, 31], projections [34], piecewise polynomial estimators [29, 30, 37, 39, 40] or linear programming estimators [27]. Here, we take profit of the properties of the RCTM to propose a new frontier estimator. Its asymptotic normality is proved and its finite sample properties are compared to other recent frontier estimators both on simulated and real datasets.

The paper is organized as follows. The definition of the RCTM and its links with classical risk measures are recalled in Section 2. A nonparametric estimator is introduced. Asymptotic properties are established in Section 3 and two applications are detailed in Section 4. The efficiency of our estimators is then illustrated on simulated data in Section 5 while Section 6 provides a motivating example in reliability. Proofs are postponed to the Appendix.

2 Regression risk measures

Let Y be a positive random variable and $X \in \mathbb{R}^p$ a random vector of regressors recorded simultaneously with Y . Assuming that (X, Y) is absolutely continuous with respect to Lebesgue measure, the probability density functions (p.d.f.) of X and Y given $X = x$ are denoted respectively by $g(\cdot)$ and $f(\cdot|x)$. For any $x \in \mathbb{R}^p$ such that $g(x) \neq 0$, the conditional distribution of Y given $X = x$ is characterized by the conditional survival function $\bar{F}(\cdot|x) = \mathbb{P}(Y > \cdot | X = x)$ or, equivalently, by the conditional quantile defined for $\alpha \in (0, 1)$ by $\bar{F}^{\leftarrow}(\alpha|x) = \inf\{t, \bar{F}(t|x) \leq \alpha\}$. In a risk analysis perspective, Y represents a loss while the conditional quantile is referred to the Regression Value at Risk and is denoted by $\text{RVaR}(\cdot|x) := \bar{F}^{\leftarrow}(\cdot|x)$. We shall also denote by $y^*(x) := \text{RVaR}(0|x) \in (0, +\infty]$ the right endpoint of Y given $X = x$. The Regression Conditional Tail Moment of level $\alpha \in (0, 1)$ and order $b \geq 0$ has been introduced in [16] and is defined by

$$\text{RCTM}_b(\alpha|x) := \mathbb{E} \left(Y^b | Y > \text{RVaR}(\alpha|x), X = x \right).$$

Let us note that this quantity may not exist for all $b \geq 0$, depending on the tail heaviness of Y given $X = x$, see Section 3 for sufficient conditions. Thanks to the RCTM tool, several risk measures have been adapted to the conditional framework: the Conditional Tail Expectation, the Conditional Tail Variance and the Conditional Tail Skewness. More specifically, the following regression risk measures are considered: the Regression Conditional Tail Expectation defined by

$$\text{RCTE}(\alpha|x) = \mathbb{E}(Y|Y > \text{RVaR}(\alpha|x), X = x) = \text{RCTM}_1(\alpha|x), \quad (1)$$

measuring the mean of losses above the RVaR, the Regression Conditional Tail Variance

$$\text{RCTV}(\alpha|x) = \mathbb{E}([Y - \text{RCTE}(\alpha|x)]^2|Y > \text{RVaR}(\alpha|x), X = x) = \text{RCTM}_2(\alpha|x) - \text{RCTM}_1^2(\alpha|x), \quad (2)$$

measuring the variability of the losses above the RVaR and the Regression Conditional Tail Skewness given by

$$\text{RCTS}(\alpha|x) = \frac{\text{RCTM}_3(\alpha|x)}{[\text{RCTV}(\alpha|x)]^{3/2}}, \quad (3)$$

assessing the asymmetry of the losses above the RVaR.

Starting from n independent copies $(X_1, Y_1), \dots, (X_n, Y_n)$ of the random vector (X, Y) , the estimation of the previous regression risk measures has been addressed in [16] when the distribution of Y given $X = x$ is heavy-tailed and for extreme levels α *i.e.* $\alpha = \alpha_n \rightarrow 0$ as $n \rightarrow \infty$. In view of (1)–(3), the main step is to estimate the RCTM, the other regression risk measures can then be estimated by a plug-in technique. To this end, remark that the RCTM can be rewritten as

$$\text{RCTM}_b(\alpha_n|x) = \frac{1}{\alpha_n} \mathbb{E}(Y^b \mathbb{I}\{Y > \text{RVaR}(\alpha_n|x)\} | X = x) =: \frac{1}{\alpha_n} \varphi_b(\text{RVaR}(\alpha_n|x)|x),$$

where $\mathbb{I}\{\cdot\}$ is the indicator function. The considered estimator of the Regression Conditional Tail Moment of level α_n and order b is thus given by the following three quantities:

$$\widehat{\text{RCTM}}_{b,n}(\alpha_n|x) := \frac{\sum_{i=1}^n \mathcal{K}_{h_n}(x - X_i) Y_i^b \mathbb{I}\{Y_i > \widehat{\text{RVaR}}_n(\alpha_n|x)\}}{\sum_{i=1}^n \alpha_n \mathcal{K}_{h_n}(x - X_i)}, \quad (4)$$

$$\widehat{\text{RVaR}}_n(\alpha_n|x) := \inf\{t, \hat{F}_n(t|x) \leq \alpha_n\}, \quad (5)$$

$$\hat{F}_n(y|x) := \frac{\sum_{i=1}^n \mathcal{K}_{k_n}(x - X_i) \mathbb{I}\{Y_i > y\}}{\sum_{i=1}^n \mathcal{K}_{k_n}(x - X_i)}, \quad (6)$$

with $\mathcal{K}_z(\cdot) := z^{-p} \mathcal{K}(\cdot/z)$, for all $z > 0$ and where $\mathcal{K}(\cdot)$ is a density on \mathbb{R}^p referred to as a kernel. Sequences (h_n) and (k_n) control the smoothness of the estimators. For the sake of simplicity, in what follows, the dependence on n for these two sequences is omitted and we let $\underline{\ell} := \min(h, k)$, $\bar{\ell} := \max(h, k)$. Observe that (4)–(6) are classical kernel estimators (see for instance [41, 43]) of conditional expectations, quantiles and survival functions. However, their use in an extreme context ($\alpha_n \rightarrow 0$ as $n \rightarrow \infty$) induces unusual difficulties in the nonparametric estimation, see the next section.

3 Main results

To derive the asymptotic properties of RCTM and of its estimator, an assumption on the right tail behavior of the conditional distribution of Y given $X = x$ is required. Since (X, Y) is supposed to be absolutely continuous with respect to Lebesgue measure, the function $\text{RVaR}(\cdot|x)$ is differentiable almost everywhere and we assume that

(A.1) There exists $\gamma(x) \in \mathbb{R}$ such that

$$\lim_{\alpha \rightarrow 0} \frac{\text{RVaR}'(t\alpha|x)}{\text{RVaR}'(\alpha|x)} = t^{-(\gamma(x)+1)},$$

locally uniformly in $t \in (0, \infty)$.

Condition (A.1) amounts to supposing that $-\text{RVaR}'(\cdot|x)$ is regularly varying at the origin with index $-(\gamma(x) + 1)$, see [6] for more details on regular variation theory. From [28, Corollary 1.1.10, equation (1.1.33)], condition (A.1) entails that the conditional distribution of Y given $X = x$ is in the maximum domain of attraction of the extreme-value distribution with extreme-value index $\gamma(x)$. The unknown function $\gamma(\cdot)$ is referred to as the conditional extreme-value index function. Let us point out that in [16], only the case $\gamma(x) > 0$ was considered, which corresponds to the situation where Y given $X = x$ has an heavy right tail (Fréchet maximum domain of attraction). Here, no assumption is made on the sign of $\gamma(x)$, and we let $\gamma_+(x) := \max(\gamma(x), 0)$. Finally, the sign of the function $\gamma(\cdot)$ in (A.1) is not supposed to be constant on the support of X , but it will appear that it should remain constant in a neighbourhood of the estimation point. To be more specific, let us consider three examples of distributions with different tail behaviors.

Example 1. (Fréchet maximum domain of attraction, $\gamma(x) > 0$). The Pareto distribution with cumulative distribution function (c.d.f.) $F(y|x) = 1 - y^{-\theta(x)}$, $y \geq 1$ and $\theta(\cdot) > 0$ verifies (A.1), the extreme-value index is $\gamma(x) = 1/\theta(x) > 0$. Note that $\bar{F}(\cdot|x)$ is regularly varying at infinity, this is the framework of [16].

Example 2. (Gumbel maximum domain of attraction, $\gamma(x) = 0$). The exponential distribution with c.d.f. $F(y|x) = 1 - \exp(-y/\theta(x))$, $y \geq 0$ and $\theta(\cdot) > 0$ verifies (A.1), the extreme-value index is $\gamma(x) = 0$. Note that $\bar{F}(\cdot|x)$ is *not* regularly varying at infinity.

Example 3. (Weibull maximum domain of attraction, $\gamma(x) < 0$). Let $y^*(\cdot)$ and $\theta(\cdot)$ be two positive functions. The considered c.d.f. is

$$F(y|x) = 1 - (1 - y/y^*(x))^{\theta(x)}, \quad \forall y \in [0, y^*(x)] \quad (7)$$

and thus

$$\text{RVaR}(\alpha|x) = y^*(x)(1 - \alpha^{1/\theta(x)}) \quad (8)$$

is differentiable with respect to $\alpha \in (0, 1]$. If, moreover, $\theta(x) < 1$, then the differentiability holds on the whole $[0, 1]$ interval. In any case, (A.1) is verified, the extreme-value index is $\gamma(x) = -1/\theta(x) < 0$ while $y^*(\cdot)$ is the frontier. The estimation of $y^*(\cdot)$ is illustrated on this particular example in Section 5.

Our first result establishes some asymptotic properties of the RCTM.

Proposition 1 (i) Suppose $y^*(x) < \infty$. Then, for all $b > 0$, $\text{RCTM}_b(\alpha|x) \rightarrow [y^*(x)]^b$ as $\alpha \rightarrow 0$.

(ii) Under **(A.1)**, for all $b \geq 0$ such that $b\gamma_+(x) < 1$,

$$\lim_{\alpha \rightarrow 0} \frac{\text{RCTM}_b(\alpha|x)}{[\text{RVaR}(\alpha|x)]^b} = \frac{1}{1 - b\gamma_+(x)}, \quad (9)$$

and $\text{RCTM}_b(\cdot|x)$ is regularly varying with index $-b\gamma_+(x)$.

First, let us highlight that Proposition 1 is an extension of the result established in [33] for the Conditional Tail Expectation ($b = 1$) in the framework of unconditional ($\gamma(x) = \gamma$) heavy-tailed ($\gamma > 0$) distributions. When $y^*(x)$ is finite, then the function $x \mapsto y^*(x)$ is called a frontier, and, from classical results of extreme-value theory, necessarily $\gamma(x) \leq 0$. In such a case, Proposition 1(i) shows that $\text{RCTM}_b(\alpha|x) \rightarrow (y^*(x))^b$ as $\alpha \rightarrow 0$ without further assumption on the distribution tail. This will be the starting point in Section 4.2 for designing a new frontier estimator. Basing on Proposition 1(ii), the asymptotic properties of RCTE, RCTV and RCTS can easily be derived and will reveal useful in Section 4.1.

Corollary 1 Assume **(A.1)** holds.

(i) If $\gamma(x) < 1$ then

$$\lim_{\alpha \rightarrow 0} \frac{\text{RCTE}(\alpha|x)}{\text{RVaR}(\alpha|x)} = \frac{1}{1 - \gamma_+(x)}.$$

(ii) If $\gamma(x) < 1/2$ then

$$\lim_{\alpha \rightarrow 0} \frac{\text{RCTV}(\alpha|x)}{[\text{RVaR}(\alpha|x)]^2} = \frac{\gamma_+^2(x)}{(1 - \gamma_+(x))^2(1 - 2\gamma_+(x))} =: \rho_1(\gamma_+(x)).$$

(iii) If $\gamma(x) < 1/3$ then

$$\lim_{\alpha \rightarrow 0} \text{RCTS}(\alpha|x) = \frac{(1 - \gamma_+(x))^3(1 - 2\gamma_+(x))^{3/2}}{\gamma_+^3(x)(1 - 3\gamma_+(x))} =: \rho_2(\gamma_+(x)).$$

(iv) If $\gamma(x) < 1/3$ then

$$\lim_{\alpha \rightarrow 0} \frac{\text{RCTV}(\alpha|x)[\text{RCTS}(\alpha|x)]^{2/3}}{[\text{RVaR}(\alpha|x)]^2} = (1 - 3\gamma_+(x))^{-2/3}.$$

It is interesting to note that, from (i), when $\gamma(x) \leq 0$, both risk measures RCTE and RVaR are asymptotically equivalent. Besides, a close study of the function ρ_2 appearing in (iii) shows that the RCTS tends to infinity when $\gamma_+(x)$ approaches 0 or $1/3$ and is asymptotically minimum for $\gamma(x) = \gamma_0$ where $\gamma_0 \simeq 0.2873$ is the unique root of equation $\gamma_0^3 + 5\gamma_0^2 - 5\gamma_0 + 1 = 0$ on $[1/4, 1/2]$. Such an extreme-value index γ_0 defines the distribution tail whose losses have the minimum asymmetry.

The asymptotic normality of (4)–(6) is obtained under additional assumptions. First, a Lipschitz condition on the p.d.f. of X is required. For all $(x, x') \in \mathbb{R}^p \times \mathbb{R}^p$, let us denote by $d(x, x')$ a distance between x and x' .

(A.2) There exists a constant $c_g > 0$ such that $|g(x) - g(x')| \leq c_g d(x, x')$.

The next assumption is devoted to the kernel function $\mathcal{K}(\cdot)$.

(A.3) $\mathcal{K}(\cdot)$ is a bounded density on \mathbb{R}^p , with support S included in the unit ball of \mathbb{R}^p .

For $\xi > 0$, the largest oscillation at point $(x, y) \in \mathbb{R}^p \times \mathbb{R}_+$ associated with the Regression Conditional Tail Moment of order $b \geq 0$ such that $b\gamma(x) < 1$ is given by

$$\omega(x, \alpha, b, \xi, h) = \sup \left\{ \left| \frac{\varphi_b(\varphi_b^\leftarrow(\beta|x)|x')}{\beta} - 1 \right| \text{ with } \left| \frac{\beta}{\alpha} - \text{RCTM}_b(\alpha|x) \right| \leq \xi \text{ and } x' \in B(x, h) \right\},$$

recalling that $\varphi_b(\cdot|x) = \bar{F}(\cdot|x)\text{RCTM}_b(\bar{F}(\cdot|x)|x)$ and where $B(x, h)$ denotes the ball centred at x with radius h . Finally, for all finite set E , let $\mathcal{L}(E) = \{e_i + e_j, (e_i, e_j) \in E \times E\} \cup E$. The first theorem establishes the asymptotic normality of the RVaR estimator defined by (5) and (6).

Theorem 1 Suppose **(A.1)**, **(A.2)** and **(A.3)** hold. Let $x \in \mathbb{R}^p$ such that $g(x) > 0$ and consider $\alpha_n \rightarrow 0$ such that $nk^p\alpha_n \rightarrow \infty$ as $n \rightarrow \infty$. If there exists $\xi > 0$ such that

$$nk^p\alpha_n (k \vee \omega(x, \alpha_n, 0, \xi, k))^2 \rightarrow 0, \quad (10)$$

then

$$(nk^p\alpha_n^{-1})^{1/2} f(\text{RVaR}(\alpha_n|x)|x) \left(\widehat{\text{RVaR}}_n(\alpha_n|x) - \text{RVaR}(\alpha_n|x) \right) \xrightarrow{d} \mathcal{N} \left(0, \frac{\|\mathcal{K}\|_2^2}{g(x)} \right).$$

This result was first established in [11, Theorem 1] but under a stronger assumption on the tail of Y given $X = x$. The Von-Mises condition used in [11] requires the twice differentiability of F . Here, it is replaced by **(A.1)** which only involves the first derivative of F . As observed in [11], condition $nk^p\alpha_n \rightarrow \infty$ limits the range of extreme Regression Value at Risk that can be estimated with a kernel method. Condition (10) implies that the bias introduced by the oscillation of the survival function

$$\omega(x, \alpha_n, 0, \xi, k) = \sup \left\{ \left| \frac{\bar{F}(\bar{F}^\leftarrow(\beta|x)|x')}{\beta} - 1 \right| \text{ with } \left| \frac{\beta}{\alpha_n} - 1 \right| \leq \xi \text{ and } x' \in B(x, k) \right\},$$

should be negligible compared to the standard-deviation of the estimator. The smaller the oscillation is, the better the nonparametric estimation procedure will perform. Moreover, (10) entails that $\omega(x, \alpha_n, 0, \xi, k) \rightarrow 0$ as $n \rightarrow \infty$. This condition can also be found in [11], it is verified under smoothness assumptions on the conditional survival function. In particular, it requires the endpoint $y^*(\cdot)$ to stay either finite or infinite in a neighbourhood of x . Similarly, the sign of the extreme-value index $\gamma(\cdot)$ should remain constant in a neighbourhood of x .

Theorem 2 Let $J \in \mathbb{N} \setminus \{0\}$ and $E := \{b_1, \dots, b_J\}$ where $b_j > 0$ for all $j = 1, \dots, J$. Suppose **(A.1)**, **(A.2)** and **(A.3)** hold and let $\alpha_n \rightarrow 0$ be a sequence satisfying $n\bar{\ell}^p\alpha_n \rightarrow \infty$ as $n \rightarrow \infty$. Let $x \in \mathbb{R}^p$ such that $g(x) > 0$ and $\gamma(x) < 1/(2b_j)$ for all $j = 1, \dots, J$. If there exists $\xi > 0$ such that

$$n\bar{\ell}^p\alpha_n \left(\bar{\ell} \vee \max_{b \in \mathcal{L}(E)} \omega(x, \alpha_n, b, \xi, \bar{\ell}) \right)^2 \rightarrow 0,$$

and $\bar{\ell}/\underline{\ell} \rightarrow 0$, then the random vector

$$(n\bar{\ell}^p\alpha_n)^{1/2} \left\{ \left(\frac{\widehat{\text{RCTM}}_{b_j, n}(\alpha_n|x)}{\text{RCTM}_{b_j}(\alpha_n|x)} - 1 \right) \right\}_{j \in \{1, \dots, J\}}$$

is asymptotically Gaussian, centred, with a $J \times J$ covariance matrix given by either $\|\mathcal{K}\|_2^2 \Sigma_E^{(1)}(x)/g(x)$ if $h/k \rightarrow 0$, or $\|\mathcal{K}\|_2^2 \Sigma_E^{(2)}(x)/g(x)$ if $k/h \rightarrow 0$, where for $(i, j) \in \{1, \dots, J\}^2$,

$$[\Sigma_E^{(1)}(x)]_{i,j} = \frac{(1 - b_i \gamma_+(x))(1 - b_j \gamma_+(x))}{1 - (b_i + b_j) \gamma_+(x)} \text{ and } [\Sigma_E^{(2)}(x)]_{i,j} = (1 - b_i \gamma_+(x))(1 - b_j \gamma_+(x)).$$

Two cases appear:

– If $\gamma(x) \leq 0$, then the asymptotic covariance matrices do not depend on $\{b_1, \dots, b_J\}$ and thus the estimators $\widehat{\text{RCTM}}_{b_j, n}(\alpha_n|x)$, $j = 1, \dots, J$ share the same rate of convergence.

– Conversely, when $\gamma(x) > 0$, the asymptotic variances are increasing functions of the RCTM order. Moreover, note that $[\Sigma_E^{(1)}(x)]_{j,j} > [\Sigma_E^{(2)}(x)]_{j,j}$ for all $j \in \{1, \dots, J\}$ and thus $k/h \rightarrow 0$ leads to more efficient estimators than $h/k \rightarrow 0$. Let us also note that, in this situation, the case $h = k$ has been investigated in [16, Theorem 1] where it was shown that diagonal terms of the covariance matrix are

$$[\Sigma_E^{(3)}(x)]_{j,j} = \frac{2b_j^2 \gamma^2(x)(1 - b_j \gamma(x))}{1 - 2b_j \gamma(x)}.$$

Routine calculations show that $[\Sigma_E^{(1)}(x)]_{j,j} > [\Sigma_E^{(3)}(x)]_{j,j}$ for all $j \in \{1, \dots, J\}$ and thus $k = h$ leads to more efficient estimators than $h/k \rightarrow 0$. The comparison between the choices $k/h \rightarrow 0$ and $h = k$ is less straightforward: $[\Sigma_E^{(2)}(x)]_{j,j} < [\Sigma_E^{(3)}(x)]_{j,j}$ for all $j \in \{1, \dots, J\}$ if and only if $b_j \gamma(x) > 1/3$.

These conclusions are however only of theoretical interest, since, in practice, $\gamma(x)$ is unknown and the bandwidths h and k have to be determined for a fixed value of the sample size n . A data-driven procedure is proposed in Section 5 but, before that, two illustrations of the above results are proposed.

4 Applications

In Paragraph 4.1, an estimation procedure is introduced for estimating the regression risk measures (1)–(3) and the associated asymptotic properties are established. These results are derived whatever the sign of the conditional extreme-value index $\gamma(x)$ is, in contrast to [16, Corollary 1, 2] which hold only under the assumption $\gamma(x) > 0$. In Paragraph 4.2, we focus on the situation where $\gamma(x) \leq 0$ and more precisely when the distribution of $Y|X = x$ is upper bounded. A new estimator of the endpoint (or equivalently the frontier) is then proposed basing on the RCTM estimator and its asymptotic normality is proved.

4.1 Estimation of extreme regression risk measures

All the regression risk measures (1)–(3) can be estimated by plugging-in the RCTM estimator defined by (4)–(6). The obtained estimators are denoted by $\widehat{\text{RCTE}}_n(\alpha_n|x)$, $\widehat{\text{RCTV}}_n(\alpha_n|x)$ and $\widehat{\text{RCTS}}_n(\alpha_n|x)$. The following corollary establishes their asymptotic normality while their asymptotic variances are given in Table 2.

Corollary 2 (i) *Under the assumptions of Theorem 2 with $E = \{1\}$ (implying $\gamma(x) < 1/2$),*

$$(n\ell^p \alpha_n)^{1/2} \left(\frac{\widehat{\text{RCTE}}_n(\alpha_n|x)}{\text{RCTE}(\alpha_n|x)} - 1 \right),$$

is asymptotically Gaussian, centred with variance $\vartheta_{\text{RCTE},1}(\gamma_+(x))\|\mathcal{K}\|_2^2/g(x)$ if $h/k \rightarrow 0$ or $\vartheta_{\text{RCTE},2}(\gamma_+(x))\|\mathcal{K}\|_2^2/g(x)$ if $k/h \rightarrow 0$.

(ii) Under the assumptions of Theorem 2 with $E = \{1, 2\}$ (implying $\gamma(x) < 1/4$),

$$(n\underline{\ell}^p \alpha_n)^{1/2} \frac{\text{RCTV}(\alpha_n|x)}{[\text{RVaR}(\alpha_n|x)]^2} \left(\frac{\widehat{\text{RCTV}}_n(\alpha_n|x)}{\text{RCTV}(\alpha_n|x)} - 1 \right),$$

is asymptotically Gaussian, centred with variance $\vartheta_{\text{RCTV},1}(\gamma_+(x))\|\mathcal{K}\|_2^2/g(x)$ if $h/k \rightarrow 0$ or $\vartheta_{\text{RCTV},2}(\gamma_+(x))\|\mathcal{K}\|_2^2/g(x)$ if $k/h \rightarrow 0$.

(iii) Under the assumptions of Theorem 2 with $E = \{1, 2, 3\}$ (implying $\gamma(x) < 1/6$), if $(n\underline{\ell}^p \alpha_n)^{1/2}[\text{RCTS}(\alpha_n|x)]^{-2/3} \rightarrow \infty$, then

$$(n\underline{\ell}^p \alpha_n)^{1/2} [\text{RCTS}(\alpha_n|x)]^{-2/3} \left(\frac{\widehat{\text{RCTS}}_n(\alpha_n|x)}{\text{RCTS}(\alpha_n|x)} - 1 \right),$$

is asymptotically Gaussian, centred with variance $\vartheta_{\text{RCTS},1}(\gamma_+(x))\|\mathcal{K}\|_2^2/g(x)$ if $h/k \rightarrow 0$ or $\vartheta_{\text{RCTS},2}(\gamma_+(x))\|\mathcal{K}\|_2^2/g(x)$ if $k/h \rightarrow 0$.

The following comments can be made:

- In the case $\gamma(x) \leq 0$, estimators of the RCTV and RCTS both have the same rate of convergence (Corollary 1(iv)) while the estimator of the RCTE converges faster (Corollary 1(i)–(iii)).
- In the case $\gamma(x) > 0$, from Corollary 1(ii) and (iii), all the previous risk measures share the same rate of convergence $(n\underline{\ell}^p \alpha_n)^{1/2}$. The asymptotic variances of

$$(n\underline{\ell}^p \alpha_n)^{1/2} \left(\frac{\widehat{\text{RCTE}}_n(\alpha_n|x)}{\text{RCTE}(\alpha_n|x)} - 1 \right), (n\underline{\ell}^p \alpha_n)^{1/2} \left(\frac{\widehat{\text{RCTV}}_n(\alpha_n|x)}{\text{RCTV}(\alpha_n|x)} - 1 \right), (n\underline{\ell}^p \alpha_n)^{1/2} \left(\frac{\widehat{\text{RCTS}}_n(\alpha_n|x)}{\text{RCTS}(\alpha_n|x)} - 1 \right),$$

are respectively given by $\tilde{\vartheta}_{\text{RCTE},\bullet}(\gamma_+(x))$, $\tilde{\vartheta}_{\text{RCTV},\bullet}(\gamma_+(x))$, $\tilde{\vartheta}_{\text{RCTS},\bullet}(\gamma_+(x))$ where

$$\begin{aligned} \tilde{\vartheta}_{\text{RCTE},\bullet}(z) &:= \vartheta_{\text{RCTE},\bullet}(z), \\ \tilde{\vartheta}_{\text{RCTV},\bullet}(z) &:= \vartheta_{\text{RCTV},\bullet}(z)/\rho_1^2(z), \\ \tilde{\vartheta}_{\text{RCTS},\bullet}(z) &:= \vartheta_{\text{RCTS},\bullet}(z)\rho_2^{4/3}(z), \end{aligned}$$

see Table 2 for details and Figure 1 for an illustration. It appears that the asymptotic variance associated with the situation $k/h \rightarrow 0$ is smaller than the asymptotic variance associated with the situation $k/h \rightarrow 0$ for all three considered estimators, even though they almost coincide for RCTS. This is consistent with the conclusions derived from Theorem 2 where it has already been observed that the case $k/h \rightarrow 0$ was the most favorable to estimate the RCTM. Since $\tilde{\vartheta}_{\text{RCTE},2}$, $\tilde{\vartheta}_{\text{RCTV},2}$ and $\tilde{\vartheta}_{\text{RCTS},2}$ are decreasing functions, the following bounds can readily be established:

$$\begin{aligned} \tilde{\vartheta}_{\text{RCTE},2}(z) &\geq 1/4 \quad \text{for all } z \in [0, 1/2], \\ \tilde{\vartheta}_{\text{RCTV},2}(z) &\geq 225/4 \quad \text{for all } z \in [0, 1/4], \\ \tilde{\vartheta}_{\text{RCTS},2}(z) &\geq 5041/4 \quad \text{for all } z \in [0, 1/6], \end{aligned}$$

and therefore the estimation of RCTV and RCTS is very unstable whatever the tail heaviness is.

4.2 Frontier estimation

This paragraph is dedicated to the estimation of the right (positive) endpoint $y^*(x) := \text{RVaR}(0|x)$ of the distribution of Y given $X = x$ in the situation where $y^*(x) < \infty$ (and thus when $\gamma(x) \leq 0$). From Proposition 1(i), the Regression Conditional Tail Moment of order $b \geq 0$ exists and

$$\text{RCTM}_b(\alpha|x) \rightarrow [y^*(x)]^b \text{ as } \alpha \rightarrow 0. \quad (11)$$

For all $b > 0$, a natural estimator of the right endpoint (or frontier) is thus given by

$$\hat{y}_{b,n}^*(x) := \left[\widehat{\text{RCTM}}_{b,n}(\alpha_n|x) \right]^{1/b}, \quad (12)$$

where α_n is a sequence converging to 0 as $n \rightarrow \infty$. In the unconditional situation, the estimation of the endpoint of a distribution has been widely studied in the extreme-value literature, see [28, Section 4.5] for an overview or [1, 19] for applications. The methods rely on the extrapolation beyond an extreme quantile via the estimation of the extreme-value index. In our situation, the adaptation of such techniques would require the estimation of the conditional extreme-value index $\gamma(x)$ which would induce additional difficulties. Our idea is to rely on the definition of the RCTM itself, which ensures that $[\text{RCTM}_b(\alpha_n|x)]^{1/b} \geq \text{RVaR}(\alpha_n|x)$. There is thus some hope that $\hat{y}_{b,n}^*(x)$ extrapolate beyond the extreme conditional quantile $\text{RVaR}(\alpha_n|x)$ without estimating $\gamma(x)$.

Let us highlight that, however, for a fixed value of n , one does not necessarily have $\hat{y}_{b,n}^*(X_i) \geq Y_i$ for all $i = 1, \dots, n$. More generally, this is also the case for robust estimators of the frontier [2, 7, 13]. For instance, in [7], an expected frontier of order m is defined. The expected frontier converges to the true frontier as $m \rightarrow \infty$ (see [7, Theorem 2.3]) similarly to (11) but it does not necessary envelop all the data points either. This property illustrates the fact that such estimators are less sensitive to extreme values or outliers than classical nonparametrical ones.

Assumption **(A.1)** is not required to justify the expression of the right endpoint estimator but it will reveal necessary to establish its asymptotic normality. Before stating this result, some notations are required. Let $a(u|x) := \bar{F}(u|x)/f(u|x)$. Under **(A.1)** with $y^*(x) < \infty$, Lemma 1, equation (19) shows that

$$\Delta_a(u|x) := \frac{a(u|x)}{u} - \gamma_+(x) = \frac{a(u|x)}{u} \rightarrow 0 \quad (13)$$

as $u \uparrow y^*(x)$.

Corollary 3 *Suppose the assumptions of Theorem 2 hold with $E = \{b\}$, $b > 0$.*

(i) *If $y^*(x) < \infty$, $|\Delta_a(\cdot|x)|$ is asymptotically decreasing and $(n\ell^p\alpha_n)^{1/2}\Delta_a(\text{RVaR}(\alpha_n|x)|x) \rightarrow 0$ then*

$$(n\ell^p\alpha_n)^{1/2} (\hat{y}_{b,n}^*(x) - \text{RVaR}(\alpha_n|x)) \xrightarrow{d} \mathcal{N}(0, \|\mathcal{K}\|_2^2(y^*(x))^2/(b^2g(x))).$$

(ii) *If, moreover, $\gamma(x) < 0$ then*

$$(n\ell^p\alpha_n)^{1/2} (\hat{y}_{b,n}^*(x) - y^*(x)) \xrightarrow{d} \mathcal{N}(0, \|\mathcal{K}\|_2^2(y^*(x))^2/(b^2g(x))).$$

Part (i) of the result states the asymptotic normality of the estimator $\hat{y}_{b,n}^*(x)$ when centred on $\text{RVaR}(\alpha_n|x)$. It is well-known that the RVaR converges to the endpoint, but centering the asymptotic distribution on the endpoint requires the additional condition

$$(n\ell^p\alpha_n)^{1/2}(\text{RVaR}(\alpha_n|x) - y^*(x)) \rightarrow 0 \quad (14)$$

as $n \rightarrow \infty$. However, in the case (ii) where $\gamma(x) < 0$, condition (14) is automatically fulfilled. To be more specific, let us consider the situation where the c.d.f. of Y given $X = x$ is defined by (7) in Example 3. From (13), it is easy to check that $\Delta_a(\cdot|x)$ is a decreasing function given by

$$\Delta_a(u|x) = \frac{1}{\theta(x)} \left(\frac{y^*(x)}{u} - 1 \right).$$

As expected, condition (14) and condition $(n\ell^p\alpha_n)^{1/2}\Delta_a(\text{RVaR}(\alpha_n|x)|x) \rightarrow 0$ are thus equivalent. Since $\text{RVaR}(\alpha|x)$ is given by (8), these conditions simply reduce to $(n\ell^p\alpha_n)^{1/2}\alpha_n^{1/\theta(x)} \rightarrow 0$. It appears that the rate of convergence $(n\ell^p\alpha_n)^{1/2}$ of the estimator is slightly smaller than $\alpha_n^{-1/\theta(x)} = \alpha_n^{\gamma(x)}$. It directly depends on the tail behavior at the endpoint. The heavier the tail is, *i.e.* the larger the extreme-value index is, the faster the convergence.

Finally, under the assumptions of Corollary 3(ii), an asymptotic confidence interval of level τ can then be established:

$$\left[\hat{y}_{b,n}^*(x) \left(1 - \frac{u_\tau}{(n\ell^p\alpha_n)^{1/2}} \frac{\|\mathcal{K}\|_2}{b\hat{g}_n(x)^{1/2}} \right), \hat{y}_{b,n}^*(x) \left(1 + \frac{u_\tau}{(n\ell^p\alpha_n)^{1/2}} \frac{\|\mathcal{K}\|_2}{b\hat{g}_n(x)^{1/2}} \right) \right], \quad (15)$$

where $\hat{g}_n(\cdot)$ is the kernel estimator of the p.d.f. of X :

$$\hat{g}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathcal{K}_h(x - X_i) \quad (16)$$

and u_τ is the $(\tau/2)$ th quantile from the survival function of the standard Gaussian distribution.

5 Validation on simulations

The performance of the frontier estimator (12) is illustrated on simulated data and compared to some other recent propositions [24, 26]. To this end, the simulation framework of [24] is used: X is a one-dimensional standard uniform random variable and the c.d.f. of Y given $X = x$ is defined by (7), see Example 3. The chosen frontier function is

$$y^*(x) = \left(\frac{1}{10} + \sin(\pi x) \right) \left[\frac{11}{10} - \frac{1}{2} \exp \left(-64 \left(x - \frac{1}{2} \right)^2 \right) \right],$$

see Figure 3 for an illustration. The shape of the unknown function is challenging to estimate since it involves large derivatives as well as both concave and convex parts. Two (positive) functions $\theta(\cdot)$ are considered: $\theta_1(x) = 1.25$ and $\theta_2(x) = 1.25 + |\cos(4\pi x)|$ for all $x \in [0, 1]$.

Selection of the hyper-parameters. The bi-quadratic kernel $\mathcal{K}(x) = 15/16(1-x^2)^2\mathbb{I}\{|x| \leq 1\}$ is selected and, for the sake of simplicity, we restrict ourselves to one common bandwidth $h = k$. Consequently, estimators (4)–(6) depend on two hyper-parameters h and α . The choice of the bandwidth h , which controls the degree of smoothing, is a recurrent problem in non-parametric statistics. Besides, the choice of α is crucial, it is equivalent to the choice of the number of upper order statistics in the non-conditional extreme-value theory. In the following, a data-driven strategy is used to select simultaneously h and α . Two sets of possible equi-spaced values are introduced: $\mathcal{H} = \{h_1 \leq \dots \leq h_J\}$ where $h_1 = 0.01$, $h_J = 0.1$ and $\mathcal{A} = \{\alpha_1 \leq \dots \leq \alpha_I\}$ with $\alpha_1 = 0.01$

and $\alpha_I = 0.1$. These values ensure that there is at least one point above $\widehat{\text{RVaR}}(\alpha_j|x)$ in the ball $B(x, h_i)$ for all $x \in [0, 1]$ and $(h_j, \alpha_j) \in \mathcal{H} \times \mathcal{A}$. The cardinal of the sets \mathcal{H} and \mathcal{A} are $I = J = 11$. The proposed data-driven strategy consists in minimizing the sampled relative L^1 – error between $\widehat{\text{RTCM}}_{2,n}$ and $\widehat{\text{RVaR}}_n^2$:

$$(h_{data}, \alpha_{data}) = \arg \min_{(h_i, \alpha_j) \in \mathcal{H} \times \mathcal{A}} \frac{1}{T} \sum_{t=1}^T \left| \frac{\widehat{\text{RTCM}}_{2,n}(\alpha_j|x_t)}{\widehat{\text{RVaR}}_n^2(\alpha_j|x_t)} - 1 \right|,$$

where $x_t = t/(T+1)$ and $T = 50$. The idea motivating this criteria is that both $\widehat{\text{RTCM}}_{2,n}(\alpha|\cdot)$ and $\widehat{\text{RVaR}}_n^2(\alpha|\cdot)$ should be close to $(y^*)^2(\cdot)$ if α and h are well chosen. To assess the behavior of the selection procedure, it is compared to an oracle strategy which consists in minimizing the relative L^1 – error between an estimator $\hat{y}^*(\cdot)$ and the true frontier $y^*(\cdot)$:

$$(h_{oracle}, \alpha_{oracle}) = \arg \min_{(h_i, \alpha_j) \in \mathcal{H} \times \mathcal{A}} \frac{1}{T} \sum_{t=1}^T \left| \frac{\hat{y}^*(x_t)}{y^*(x_t)} - 1 \right|.$$

Of course, the oracle strategy cannot be used in practice to select h and α since the true function $y^*(\cdot)$ is unknown. However, it provides a lower bound on the L_1 – error that can be reached with the proposed estimators.

Competing estimators. Ten estimators $\hat{y}_{1,n}^*, \dots, \hat{y}_{10,n}^*$ deduced from (12) are compared with $\widehat{\text{RVaR}}_n$ and three estimators $\hat{y}_n^{(*,gj)}$, $\hat{y}_n^{(*,mc)}$ and $\hat{y}_n^{(*,mv)}$ from [24, 26]. All three previous estimators are based on a kernel estimator of the high order moments of Y given $X = x$:

$$\hat{\mu}_{p,n}(x) = \frac{1}{n} \sum_{i=1}^n Y_i^p \mathcal{K}_h(x - X_i).$$

where $p \rightarrow \infty$ and $h \rightarrow 0$ as $n \rightarrow \infty$. The first one relies on the assumption that Y given $X = x$ is uniformly distributed on $[0, g(x)]$:

$$\hat{y}_n^{(*,gj)}(x) = ((p+1)\hat{\mu}_{p,n}(x)/\hat{g}_n(x))^{1/p},$$

where $\hat{g}_n(\cdot)$ is the kernel estimator (16) of the density of X . Estimators $\hat{y}_n^{(*,mc)}$ and $\hat{y}_n^{(*,mv)}$ do not rely on a parametric assumption. They can be written as

$$\frac{1}{\hat{y}_n^{*,\bullet}(x)} = \frac{1}{a} \left[((a+1)p+1) \frac{\hat{\mu}_{(a+1)p,n}(x)}{\hat{\mu}_{(a+1)p+1,n}(x)} - (p+1) \frac{\hat{\mu}_{p,n}(x)}{\hat{\mu}_{p+1,n}(x)} \right],$$

where $a > 0$ is an additional parameter and $\bullet \in \{mc, mv\}$. The difference between $\hat{y}_n^{(*,mc)}$ and $\hat{y}_n^{(*,mv)}$ lies in the choice of the hyper-parameters h , p and a , see [24] for implementation details.

Results. The finite sample performance of the estimators is assessed on $N = 500$ replications of samples of size $n = 500$. The choices $(h_{data}, \alpha_{data})$ and $(h_{oracle}, \alpha_{oracle})$ associated with $\hat{y}_{1,n}^*, \dots, \hat{y}_{10,n}^*$ and $\widehat{\text{RVaR}}$ are computed on each replication with the two strategies. The minimum, maximum and mean L^1 – errors associated with $(h_{data}, \alpha_{data})$ are given in Table 3. The results associated with $\hat{y}_n^{(*,gj)}$, $\hat{y}_n^{(*,mc)}$ and $\hat{y}_n^{(*,mv)}$ are reported from [24]. It appears that $\hat{y}_{1,n}^*$ does not yield very good results but $\hat{y}_{2,n}^*, \dots, \hat{y}_{10,n}^*$ all perform better than $\widehat{\text{RVaR}}$, $\hat{y}_n^{(*,gj)}$, $\hat{y}_n^{(*,mc)}$ and

$\hat{y}_n^{(*,mv)}$ in both situations $\theta(\cdot) = \theta_1(\cdot)$ and $\theta(\cdot) = \theta_2(\cdot)$. Among them, $\hat{y}_{7,n}^*$ yields the best results but the behavior of $\hat{y}_{5,n}^*$ and $\hat{y}_{6,n}^*$ are very close. The performance of the data-driven selection of the hyper-parameters is compared to the oracle one on Figure 2. Histograms of the L^1 - errors associated with each strategies are displayed for $\hat{y}_{7,n}^*$ and for both functions $\theta_1(\cdot)$ and $\theta_2(\cdot)$. Unsurprisingly, the oracle strategy yields smaller errors than the data-driven one, but the large overlap of the histograms shows that the data-driven selection procedure yields reasonable results. Figure 3 provides the best, the worst and the median estimation of the frontier respectively corresponding to the min, max and median of the mean L^1 - errors. It appears study that $\hat{y}_{7,n}^*$ combined with the data-driven hyper-parameters selection is a reasonable frontier estimator both in terms of stability and precision.

6 Illustration on real data

As an illustration, an application to the reliability of nuclear reactors is proposed. The data consist in $n = 254$ non-irradiated representative steels obtained from the US Electric Power Research Institute. The variable of interest Y is the fracture toughness and the unidimensional covariate X is the temperature measured in degrees Fahrenheit. As the temperature decreases, the steels fissure more easily (see Figure 4). In a worst case scenario, it is important to know the upper limit of fracture toughness of each material as a function of the temperature, that is $y^*(x)$ the endpoint of Y given $X = x$. In view of Figure 4, one may assume that the frontier $y^*(x)$ is an increasing function of the covariate x . In such a case, the frontier can also be interpreted as the endpoint of Y given $X \leq x$. Introducing this prior information opens the way to specific estimation techniques, see for instance [12, 14, 17, 23]. We also refer to [2, 7, 13] for the definition of robust estimators. Here, the estimator $\hat{y}_{7,n}^*$ is compared to the spline-based estimators CS-B and QS-B recently introduced [12] for monotone boundaries (CS and QS refer respectively to cubic and quadratic splines). The BIC criterion is used to determine the complexity of the spline approximation. The hyper-parameters associated with $\hat{y}_{7,n}^*$ are chosen in the sets $\mathcal{H} = \{17, 18, \dots, 120\}$ and $\mathcal{A} = \{0.01, 0.015, \dots, 0.1\}$ using the data-driven procedure presented in Section 5. The selection yields $(h_{data}, \alpha_{data}) = (98, 0.085)$, results are depicted on Figure 4. All three estimators yield increasing functions even though this constraint was not implemented in $\hat{y}_{7,n}^*$. Also, the three estimated frontiers coincide on the range $x \in [-120, -30]$. Results are slightly different outside this interval: CS-B and QS-B estimators simply interpolate the boundary points whereas $\hat{y}_{7,n}^*$ estimates a heavier tail and thus a higher value for the limit of fracture toughness. Basing on (15), Figure 4 provides pointwise 95% asymptotic confidence intervals centered on $\hat{y}_{7,n}^*$ for the frontier. The estimation of $\gamma(\cdot)$ is also displayed on Figure 5 together with the associated pointwise 95% asymptotic confidence intervals. The following estimator, introduced in [11], has been implemented:

$$\hat{\gamma}_n^{\text{RP},1}(x) = \frac{1}{\log(2/3)} \log \left(\frac{\widehat{\text{RVaR}}_n(\alpha_n|x) - \widehat{\text{RVaR}}_n(2\alpha_n/3|x)}{\widehat{\text{RVaR}}_n(2\alpha_n/3|x) - \widehat{\text{RVaR}}_n(4\alpha_n/9|x)} \right).$$

Note that this estimator also coincides with the kernel Pickands estimator introduced and studied in [10] in the case $\gamma(x) > 0$. It appears that the estimation of $\gamma(x)$ is negative for most of the values of the covariate x and the 95% confidence interval is most of the time included in \mathbb{R}^- . It seems thus reasonable to assume that $y^*(x) < \infty$ for most of the values of the covariate x . However, the

estimations $\hat{y}_{7,n}^*(x)$ for which $\hat{\gamma}_n^{\text{RP},1}(x) \geq 0$ should be considered with great care. To conclude, we would like to stress two possible improvements:

- In this work, we do not propose an automatic procedure for selecting b in $\hat{y}_{b,n}^*$. From the practical point of view, our opinion is that any choice of $b \geq 4$ is satisfying, the estimation being very stable over this threshold. As illustrated on Figure 6, $\hat{y}_{4,n}^*$, $\hat{y}_{5,n}^*$, $\hat{y}_{6,n}^*$ and $\hat{y}_{7,n}^*$ are very close to each other. In fact, using the pointwise confidence intervals (15), at each point x , it is possible to show that these four estimators are not significantly different at the 5% level. From the theoretical point of view, Corollary 3 shows that the asymptotic variance is a decreasing function of b . Since there is no counterpart on the bias, this would suggest to choose $b = b_n \rightarrow \infty$ and to investigate the asymptotic behavior of this new estimator. This approach would be similar to the kernel regression on high order moments developed in [24].

- It is well-known that non-parametric estimators based on Parzen-Rosenblatt kernels may suffer from a lack of performance on the boundaries of the estimation interval [21]. This phenomenon appears on Figure 4. When x is large, $\hat{y}_{7,n}^*(x)$ slightly underestimates the true frontier. To overcome this limitation, symmetrization [8] and jackknife [36] techniques have been developed. They could be adapted to our framework.

References

- [1] Aarssen, K., and de Haan, L. (1994). On the maximal life span of humans. *Mathematical Population Studies*, **4**(4), 259–281.
- [2] Aragon, Y., Daouia, A., and Thomas-Agnan, C. (2005). Nonparametric frontier estimation: a conditional quantile-based approach. *Journal of Econometric Theory*, **21**(2), 358–389.
- [3] Artzner, P., Delbaen, F., Eber, J.M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**, 203–228.
- [4] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels J. (2004). *Statistics of extremes: Theory and applications*, Wiley.
- [5] Berline, A., Gannoun, A., and Matzner-Lober, E. (2001). Asymptotic normality of convergent estimates of conditional quantiles. *Statistics*, **35**, 139–169.
- [6] Bingham, N.H., Goldie, C.M., and Teugels, J.L. (1987). *Regular Variation*, Cambridge University Press.
- [7] Cazals, C., Florens, J.-P., and Simar, L. (2002). Nonparametric frontier estimation: A robust approach. *Journal of Econometrics*, **106**(1), 1–25.
- [8] Cowling, A., and Hall, P. (1996). On pseudodata methods for removing boundary effects in kernel density estimation. *Journal of the Royal Statistical Society B*, **58**, 551–563.
- [9] Daouia, A., Florens, J.-P., and Simar, L. (2010). Frontier estimation and extreme value theory. *Bernoulli*, **16**, 1039–1063.
- [10] Daouia, A., Gardes, L., Girard, S. & Lekina, A. (2011). Kernel estimators of extreme level curves, *TEST*, **20**, 311–333.

- [11] Daouia, A., Gardes, L., and Girard, S. (2013). On kernel smoothing for extremal quantile regression. *Bernoulli*, **19**, 2557–2589.
- [12] Daouia, A., Noh, H., and Park, B.U. (2016). Data envelope fitting with constrained polynomial spline, *Journal of the Royal Statistical Society, B*, **78**(1), 3–36.
- [13] Daouia, A., and Simar, L. (2005). Robust nonparametric estimators of monotone boundaries. *Journal of Multivariate Analysis*, **96**, 311–331.
- [14] Deprins, D., Simar, L., and Tulkens, H. (1984). Measuring Labor Efficiency in Post Offices. in *The Performance of Public Enterprises: Concepts and Measurements* by M. Marchand, P. Pestieau and H. Tulkens, North Holland ed, Amsterdam.
- [15] Drees, H. (1995). Refined Pickands estimators of the extreme value index. *The Annals of Statistics*, **23**, 2059–2080.
- [16] El Methni, J., Gardes, L., and Girard, S. (2014). Nonparametric estimation of extreme risk measures from conditional heavy-tailed distributions. *Scandinavian Journal of Statistics*, **41**(4), 988–1012.
- [17] Farrel, M.J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society A*, **120**, 253–281.
- [18] Fraga Alves, M. I., Gomes, M. I., de Haan, L., and Neves, C. (2007). A note on second order conditions in extreme value theory: linking general and heavy tail conditions. *REVSTAT - Statistical Journal*, **5**(3), 285–304.
- [19] Fraga Alves, M. I., Neves, C., and Rosário, P. (2017). A general estimator for the right endpoint with an application to supercentenarian women’s records. *Extremes*, **20**, 199–237.
- [20] Gardes, L., and Girard, S. (2010). Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels. *Extremes*, **13**, 177–204.
- [21] Gasser, T., and Müller, H. (1979). Kernel estimation of regression functions. In: Gasser T., Rosenblatt M. (eds) *Smoothing techniques for curve estimation*, Lecture Notes in Mathematics, vol 757, pp. 23–68, Springer, Berlin Heidelberg.
- [22] Geffroy, J. (1964). Sur un problème d’estimation géométrique. *Publications de l’Institut de Statistique de l’Université de Paris*, **XIII**, 191–200.
- [23] Gijbels, I., Mammen, E., Park, B.U., and Simar, L. (1999). On estimation of monotone and concave frontier functions. *Journal of the American Statistical Association*, **94**, 220–228.
- [24] Girard, S., Guillou, A., and Stupfler, G. (2013). Frontier estimation with kernel regression on high order moments. *Journal of Multivariate Analysis*, **116**, 172–189.
- [25] Girard, S., and Jacob, P. (2004). Extreme values and kernel estimates of point processes boundaries. *ESAIM: Probability and Statistics*, **8**, 150–168.
- [26] Girard, S., and Jacob, P. (2008). Frontier estimation via kernel regression on high power transformed data. *Journal of Multivariate Analysis*, **99**, 403–420.

- [27] Girard, S., and Nazin, A.V. (2014). L1-optimal linear programming estimator for periodic frontier functions with Holder continuous derivative. *Automation and Remote Control*, **75**(12), 2152–2169.
- [28] de Haan, L., and Ferreira, A. (2006). *Extreme Value Theory: An introduction*, Springer Series in Operations Research and Financial Engineering, Springer.
- [29] Hall, P., Park, B. U., and Stern, S. E. (1998). On polynomial estimators of frontiers and boundaries. *Journal of Multivariate Analysis*, **66**, 71–98.
- [30] Härdle, W., Hall, P., and Simar, L. (1995). Iterated bootstrap with application to frontier models. *J. Productivity Anal.*, **6**, 63–76.
- [31] Härdle, W., Park, B. U., and Tsybakov, A. B. (1995). Estimation of a non sharp support boundaries. *Journal of Multivariate Analysis*, **43**, 205–218.
- [32] Hong, J., and Elshahat, A. (2010). Conditional tail variance and conditional tail skewness. *Journal of Financial and Economic Practice*, **10**(1), 147–156.
- [33] Hua, L., and Joe, H. (2011). Second order regular variation and conditional tail expectation of multiple risks. *Insurance: Mathematics and Economics*, **49**(3), 537–546.
- [34] Jacob, P., and Suquet, P. (1995). Estimating the edge of a Poisson process by orthogonal series. *Journal of Statistical Planning and Inference*, **46**, 215–234.
- [35] Jorion, P. (2007). *Value at risk: the new benchmark for managing financial risk*, McGraw-Hill New York.
- [36] Kyung-Joon, C., and Schucany, W. (1998). Nonparametric kernel regression estimation near endpoints. *Journal of Statistical Planning and Inference*, **66**, 289–304.
- [37] Knight, K. (2001). Limiting distributions of linear programming estimators. *Extremes*, **4**(2), 87–103.
- [38] Koenker, R., and Geling, O. (2001). Reappraising medfly longevity: A quantile regression survival analysis. *Journal of American Statistical Association*, **96**, 458–468.
- [39] Korostelev, A., Simar, L., and Tsybakov, A. B. (1995). Efficient estimation of monotone boundaries. *The Annals of Statistics*, **23**, 476–489.
- [40] Korostelev, A.P., and Tsybakov, A.B. (1993). Minimax theory of image reconstruction. in *Lecture Notes in Statistics*, **82**, Springer-Verlag, New York.
- [41] Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, **33**, 1065–1076.
- [42] Resnick, S.I. (2008). *Extreme Values, Regular Variation, and Point Processes*, Springer.
- [43] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, 832–837.

- [44] Samanta, T. (1989). Non-parametric estimation of conditional quantiles. *Statistics and Probability Letters*, **7**, 407–412.
- [45] Smith, R.L. (1989). Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone (with discussion). *Statistical Science*, **4**, 367–393.
- [46] Stone, C.J. (1977). Consistent nonparametric regression (with discussion). *The Annals of Statistics*, **5**, 595–645.
- [47] Stute, W. (1986). Conditional empirical processes. *The Annals of Statistics*, **14**, 638–647.
- [48] Tsay, R.S. (2002). *Analysis of financial time series*. Wiley, New-York.
- [49] Valdez, E.A. (2005). Tail conditional variance for elliptically contoured distributions. *Belgian Actuarial Bulletin*, **5**, 26–36.

Acknowledgments.

The authors would like to thank the Associate Editor and an anonymous referee for their helpful suggestions. They led to a significant improvement of the initial manuscript.

$U(\cdot x)$	$\beta(x)$	$a_U(\cdot x)$
$\text{RVaR}(\cdot x)$ $\varphi_b^{\leftarrow}(\cdot x)$	$\gamma(x)$ $\tilde{\gamma}_b(x) := \gamma(x)/(1 - b\gamma_+(x))$	$a(\cdot x) := \bar{F}(\cdot x)/f(\cdot x)$ $\tilde{a}_b(\cdot x) := a(\cdot x)/(1 - b\gamma_+(x))$

Table 1: Index of regular variation and auxiliary function associated with $\text{RVaR}(\cdot|x)$ and $\varphi_b^{\leftarrow}(\cdot|x)$.

7 Appendix: Proofs

7.1 Preliminary results

We start with useful results on regularly varying functions. The set of regularly varying functions at 0 with index $\beta \in \mathbb{R}$ is denoted by \mathcal{RV}_β . Recall that a function $V(\cdot) \in \mathcal{RV}_\beta$ if $V(\cdot)$ is asymptotically positive and such that $V(t\alpha)/V(\alpha) \rightarrow t^\beta$ as $\alpha \rightarrow 0$, locally uniformly in $t \in (0, \infty)$.

Lemma 1 *Let $U(\cdot|x)$ be a decreasing and differentiable function with support $(0, M)$, $M > 0$. Let us introduce the positive function $a_U(\cdot|x)$ such that for $p \in (0, M)$, $a_U(U(p|x)|x) = -pU'(p|x)$ and the function $L_z(\cdot)$ defined for all $(z, u) \in \mathbb{R}^2$ by $L_z(u) := \int_1^u v^{z-1} dv$. If $U'(\cdot|x)$ is regularly varying with index $-(\beta(x) + 1)$, $\beta(x) \in \mathbb{R}$, then locally uniformly in $s \in (0, \infty)$,*

$$\lim_{p \rightarrow 0} \frac{U(sp|x) - U(p|x)}{a_U(U(p|x)|x)} = L_{\beta(x)}(1/s), \quad (17)$$

which is equivalent to

$$\lim_{p \rightarrow 0} \frac{U^{\leftarrow}(U(p|x) + ta_U(U(p|x)|x)|x)}{p} = \frac{1}{L_{\beta(x)}^{\leftarrow}(t)}, \quad (18)$$

locally uniformly in $t \in (-1/\beta_+(x), -1/\beta_-(x))$. Finally, if $U'(\cdot|x) \in \mathcal{RV}_{-(\beta(x)+1)}$, then

$$\lim_{p \rightarrow 0} \frac{a_U(U(p|x)|x)}{U(p|x)} = \beta_+(x). \quad (19)$$

The function $a_U(\cdot|x)$ is called the auxiliary function of $U(\cdot|x)$. Obviously, under **(A.1)**, the function $\text{RVaR}(\cdot|x)$ satisfies the assumptions of Lemma 1 (see the first row of Table 1).

Proof. First remark that condition $U'(\cdot|x) \in \mathcal{RV}_{-(\beta(x)+1)}$ coincides with condition (1.1.33) in [28, Corollary 1.1.10] which implies (17). Second, condition (17) is equivalent to condition (1.1.20) in [28, Theorem 1.1.6], with the auxiliary function $a_U(\cdot|x)$, which is also equivalent to (18). Finally, from [18, Lemma 3.1] condition (18) entails (19). \blacksquare

Since for any $b \geq 0$ such that the moment of order b of Y exists, $\varphi_b(\cdot|x) = \bar{F}(\cdot|x)\text{RCTM}_b(\bar{F}(\cdot|x)|x)$ or, equivalently, the Regression Conditional Tail Moment of level α and order b is given by $\text{RCTM}_b(\alpha|x) = \alpha^{-1}\varphi_b(\text{RVaR}(\alpha|x)|x)$. The next lemma is dedicated to the analysis of the function $\varphi_b(\cdot|x)$.

Lemma 2 Assume that (A.1) holds and let $b \geq 0$ such that $b\gamma_+(x) < 1$.

(i) The function $\varphi_b(\cdot|x)$ is such that

$$\lim_{y \uparrow y^*(x)} \frac{\varphi_b(y|x)}{y^b \bar{F}(y|x)} = \frac{1}{1 - b\gamma_+(x)}.$$

(ii) The function $\varphi_b(\cdot|x)$ is differentiable with derivative $\varphi'_b(\cdot|x)$ such that

$$\lim_{y \uparrow y^*(x)} \frac{\varphi'_b(y|x)}{\varphi_b(y|x)} \frac{\bar{F}(y|x)}{f(y|x)} = b\gamma_+(x) - 1.$$

Proof. (i) First, an integration by part leads to

$$\varphi_b(y|x) = \int_y^{y^*(x)} z^b F'(z|x) dz = y^b \bar{F}(y|x) + b \int_y^{y^*(x)} z^{b-1} \bar{F}(z|x) dz,$$

since, for all $b \in [0, 1/\gamma_+(x))$, $y^b \bar{F}(y|x) \rightarrow 0$ as $y \uparrow y^*(x)$, see [28, Exercise 1.11]. Straightforward calculations yield

$$\varphi_b(y|x) = (1 - b\gamma_+(x))^{-1} y^b \bar{F}(y|x) (1 + bR_b(y|x)) \quad (20)$$

with

$$R_b(y|x) = (1 - b\gamma_+(x)) \int_y^{y^*(x)} \frac{z^{b-1} \bar{F}(z|x)}{y^b \bar{F}(y|x)} dz - \gamma_+(x). \quad (21)$$

It thus remains to show that $R_b(y|x) \rightarrow 0$ as $y \uparrow y^*(x)$. Three cases are considered.

– In the case $\gamma(x) > 0$, one has $y^*(x) = \infty$ and, from [42, Theorem 0.6],

$$\lim_{y \rightarrow \infty} \int_y^\infty \frac{z^{b-1} \bar{F}(z|x)}{y^b \bar{F}(y|x)} dz = \frac{\gamma(x)}{1 - b\gamma(x)}.$$

Then, $R_b(y|x) \rightarrow 0$ as $y \uparrow y^*(x)$.

– Let us now consider the case $\gamma(x) \leq 0$ with $y^*(x) < \infty$. One has,

$$\int_y^{y^*(x)} \frac{z^{b-1} \bar{F}(z|x)}{y^b \bar{F}(y|x)} dz = \int_1^{y^*(x)/y} t^{b-1} \exp \left\{ - \int_y^{ty} \frac{f(u|x)}{\bar{F}(u|x)} du \right\} dt.$$

Under (A.1), applying Lemma 1, equation (19) with $U(\cdot) = \text{RVar}(\cdot|x)$ entails

$$\lim_{y \uparrow y^*(x)} \frac{\bar{F}(y|x)}{y f(y|x)} = \lim_{y \uparrow y^*(x)} \frac{a(y|x)}{y} = 0.$$

Hence, for $\varepsilon \in (0, 1/(2b))$, there exists $\mathcal{Y} > 0$ such that for $y > \mathcal{Y}$, $t > 1$ and $u \in [y, ty]$, $uf(u|x)/\bar{F}(u|x) > 1/\varepsilon$ leading to

$$0 \leq \int_y^{y^*(x)} \frac{z^{b-1} \bar{F}(z|x)}{y^b \bar{F}(y|x)} dz \leq \frac{\varepsilon}{1 - b\varepsilon} \leq 2\varepsilon. \quad (22)$$

– Similar calculations show that (22) also hold in the case $\gamma(x) = 0$ with $y^*(x) = \infty$. The proof of (i) is then complete since (22) is true for all $\varepsilon \in (0, 1/(2b))$.

(ii) The proof is straightforward remarking that $\varphi'_b(y|x) = -y^b f(y|x)$ and using (i) of Lemma 2. ■

The next lemma provides a control of the approximation error in Proposition 1(i) in the situation where $y^*(x) < \infty$.

Lemma 3 *Let $x \in \mathbb{R}^p$ such that $g(x) > 0$. Suppose (A.1) holds and $y^*(x) < \infty$. Introduce $\Delta_a(u|x) := a(u|x)/u$. If $|\Delta_a(\cdot|x)|$ is asymptotically decreasing then, for all $b \geq 0$,*

$$\text{RCTM}_b(\alpha|x) = [\text{RVaR}(\alpha|x)]^b [1 + O(\Delta_a(\text{RVaR}(\alpha|x)|x))],$$

as $\alpha \rightarrow 0$.

Proof. Recall that since $y^*(x) < \infty$, $\gamma(x) \leq 0$ and thus $\gamma_+(x) = 0$. As a consequence, RCTM_b exists for all $b \geq 0$ and Proposition 1(ii) states that

$$\text{RCTM}_b(\alpha|x) = [\text{RVaR}(\alpha|x)]^b (1 + o(1)).$$

We start with the following equality derived from (20) in the proof of Lemma 2(i):

$$\text{RCTM}_b(\alpha|x) = [\text{RVaR}(\alpha|x)]^b [1 + bR_b(\text{RVaR}(\alpha|x)|x)],$$

where $R_b(\cdot|x)$ is defined in (21). Since $y^*(x) < \infty$, elementary calculations yield

$$R_b(\text{RVaR}(\alpha|x)|x) = \int_1^{y^*(x)/\text{RVaR}(\alpha|x)} t^{b-1} \exp \left\{ - \int_{\text{RVaR}(\alpha|x)}^{t\text{RVaR}(\alpha|x)} \frac{1}{u} \frac{1}{\Delta_a(u|x)} du \right\} dt.$$

By assumption, $-1/\Delta_a(u|x) \leq -1/\Delta_a(\text{RVaR}(\alpha|x)|x)$ for $u \in [\text{RVaR}(\alpha|x), t\text{RVaR}(\alpha|x)]$, and thus

$$0 \leq R_b(\text{RVaR}(\alpha|x)|x) \leq \frac{1}{[\Delta_a(\text{RVaR}(\alpha|x)|x)]^{-1} - b} \left[1 - \left(\frac{y^*(x)}{\text{RVaR}(\alpha|x)} \right)^{b - [\Delta_a(\text{RVaR}(\alpha|x)|x)]^{-1}} \right].$$

Since $\Delta_a(\text{RVaR}(\alpha|x)|x) \rightarrow 0$ as $\alpha \rightarrow 0$, one has

$$0 \leq R_b(\text{RVaR}(\alpha|x)|x) \leq \Delta_a(\text{RVaR}(\alpha|x)|x) / [1 - b\Delta_a(\text{RVaR}(\alpha|x)|x)]$$

for α small enough. As a conclusion,

$$\text{RCTM}_b(\alpha|x) = [\text{RVaR}(\alpha|x)]^b [1 + O(\Delta_a(\text{RVaR}(\alpha|x)|x))]$$

and the result is proved. ■

The next lemma shows that the function $\varphi_b^{\leftarrow}(\cdot|x)$ satisfies the assumption of Lemma 1 (see Table 1).

Lemma 4 *Under (A.1), for all $b \geq 0$ such that $b\gamma_+(x) < 1$, the derivative $(\varphi_b^{\leftarrow})'(\cdot|x)$ of $\varphi_b^{\leftarrow}(\cdot|x)$ belongs to the set $\mathcal{RV}_{-(\tilde{\gamma}_b(x)+1)}$ with $\tilde{\gamma}_b(x) := \gamma(x)/(1 - b\gamma_+(x))$ and $\tilde{a}_b(\cdot|x) := a(\cdot|x)/(1 - b\gamma_+(x))$.*

Proof. For $t \in (0, \infty)$ and $\alpha \in (0, 1)$, remark that

$$\frac{(\varphi_b^{\leftarrow})'(t\alpha|x)}{(\varphi_b^{\leftarrow})'(\alpha|x)} = \frac{\varphi_b'(\varphi_b^{\leftarrow}(\alpha|x)|x)}{\varphi_b'(\varphi_b^{\leftarrow}(t\alpha|x)|x)}.$$

Recalling that $a(\cdot|x) = \bar{F}(\cdot|x)/f(\cdot|x)$, Lemma 2(ii) entails that

$$\frac{(\varphi_b^{\leftarrow})'(t\alpha|x)}{(\varphi_b^{\leftarrow})'(\alpha|x)} = \frac{1}{t} \frac{a(\varphi_b^{\leftarrow}(t\alpha|x)|x)}{a(\varphi_b^{\leftarrow}(\alpha|x)|x)} (1 + o(1)),$$

as α goes to 0. It thus remains to prove that the function $a(\varphi_b^{\leftarrow}(\cdot|x)|x)$ belongs to $\mathcal{RV}_{-\tilde{\gamma}_b(x)}$. Remark that under **(A.1)**, the function $A(\cdot|x) := a(\text{RVar}(\cdot|x)|x) = -\alpha \text{RVar}'(\alpha|x) \in \mathcal{RV}_{-\gamma(x)}$. Introduce the function $V_b(\cdot|x) := \bar{F}(\varphi_b^{\leftarrow}(\cdot|x)|x)$. Since $a(\varphi_b^{\leftarrow}(\cdot|x)|x) = A(V_b(\cdot|x)|x)$, it then suffices to prove that $V_b(\cdot|x) \in \mathcal{RV}_{(1-b\gamma_+(x))^{-1}}$. To this end, note that for all $t > 0$,

$$\frac{V_b(t\alpha|x)}{V_b(\alpha|x)} = \frac{\bar{F}(\varphi_b^{\leftarrow}(\alpha|x) + W_{b,t}(\alpha|x)a(\varphi_b^{\leftarrow}(\alpha|x)|x)|x)}{\bar{F}(\varphi_b^{\leftarrow}(\alpha|x)|x)},$$

with $W_{b,t}(\alpha|x) := (\varphi_b^{\leftarrow}(t\alpha|x) - \varphi_b^{\leftarrow}(\alpha|x))/a(\varphi_b^{\leftarrow}(\alpha|x)|x)$. Our goal is to prove that

$$\lim_{\alpha \rightarrow 0} W_{b,t}(\alpha|x) = L_{\gamma(x)} \left(t^{-(1-b\gamma_+(x))^{-1}} \right). \quad (23)$$

If (23) holds, then Lemma 1, equation (18) applied to the function $U(\cdot|x) = \text{RVar}(\cdot|x)$ with $p = \bar{F}(\varphi_b^{\leftarrow}(\alpha|x)|x)$ entails that $V_b(\cdot|x) \in \mathcal{RV}_{(1-b\gamma_+(x))^{-1}}$ which is the desired result. Straightforward calculations show that (23) is equivalent to

$$\lim_{\alpha \rightarrow 0} \frac{\varphi_b^{\leftarrow}(t\alpha|x) - \varphi_b^{\leftarrow}(\alpha|x)}{\tilde{a}_b(\varphi_b^{\leftarrow}(\alpha|x)|x)} = L_{\tilde{\gamma}_b(x)}(1/t).$$

As a consequence of the equivalence (17) \Leftrightarrow (18) in Lemma 1, limit (23) is also equivalent to

$$\lim_{y \uparrow y^*(x)} \frac{\varphi_b(y + t\tilde{a}_b(y|x)|x)}{\varphi_b(y|x)} = \frac{1}{L_{\tilde{\gamma}_b(x)}^{\leftarrow}(t)}. \quad (24)$$

Applying Lemma 1, equation (19) to the function $\text{RVar}(\cdot|x)$ leads to $a(y|x)/y \rightarrow \gamma_+(x)$ as y tends to the endpoint $y^*(x)$. Hence,

$$\lim_{y \uparrow y^*(x)} \frac{y + t\tilde{a}_b(y|x)|x}{y} = 1 + t \frac{\gamma_+(x)}{1 - b\gamma_+(x)}.$$

Consequently, $y + t\tilde{a}_b(y|x)|x \rightarrow y^*(x)$ as $y \rightarrow y^*(x)$. Thus, applying Lemma 2(i) together with Lemma 1, equation (18) leads to

$$\lim_{y \uparrow y^*(x)} \frac{\varphi_b(y + t\tilde{a}_b(y|x)|x)}{\varphi_b(y|x)} = \left(1 + t \frac{\gamma_+(x)}{1 - b\gamma_+(x)} \right)^b \frac{1}{L_{\gamma(x)}^{\leftarrow}(t/(1 - b\gamma_+(x)))}.$$

Considering separately the cases $\gamma(x) > 0$ and $\gamma(x) \leq 0$, it is easy to check that this limit is equal to $1/L_{\tilde{\gamma}_b(x)}^{\leftarrow}(t)$ which concludes the proof. \blacksquare

The Regression Conditional Tail Moment estimator defined in (4) is given by $\widehat{\text{RCTM}}_{b,n}(\alpha_n|x) = \alpha_n^{-1} \hat{\varphi}_{b,n}(\widehat{\text{RVar}}_n(\alpha_n|x)|x)$ where $\hat{\varphi}_{b,n}(\cdot|x) = \hat{\psi}_{b,n}(\cdot|x)/\hat{g}_n(x)$ with

$$\hat{\psi}_{b,n}(y|x) = \frac{1}{n} \sum_{i=1}^n \mathcal{K}_h(x - X_i) Y_i^b \mathbb{I}\{Y_i > y\}$$

an estimator of $\psi_b(y|x) = g(x)\varphi_b(y|x)$ and $\hat{g}_n(\cdot)$ the kernel estimator (16) of the p.d.f. of X . Let us finally introduce some further notations. Let $y_n(x)$ be a sequence such that $y_n(x) \uparrow y^*(x)$ and let $J \in \mathbb{N} \setminus \{0\}$. For $j = 1, \dots, J$, let $y_{n,j}(x) := y_n(x) + t_{n,j}(x)a(y_n(x)|x)$ where $t_{n,j}(x)$ are sequences converging to 0 as n goes to infinity. Recall that under **(A.1)**, $a(y|x)/y \rightarrow \gamma_+(x)$ as $y \uparrow y^*(x)$ (see Lemma 1, equation (19)). Hence, for all $j = 1, \dots, J$, $y_{n,j}(x) = y_n(x)(1 + o(1))$

as $n \rightarrow \infty$.

The next lemma is dedicated to the study of the asymptotic behavior of the random vector $\{\hat{\varphi}_{b_j,n}(y_{n,j}(x)|x)\}_{j=1,\dots,J}$ with $b_1 \geq 0, \dots, b_J \geq 0$. The oscillations of the functions $\varphi_b(y_{n,j}(x)|\cdot)$ with $b \in \mathcal{L}(E)$ where $E := \{b_1, \dots, b_J\}$ are controlled by

$$\Omega_n(x, h) := \max_{\substack{b \in \mathcal{L}(E) \\ j=1, \dots, J}} \sup \left\{ \left| \frac{\varphi_b(y_{n,j}(x)|x')}{\varphi_b(y_{n,j}(x)|x)} - 1 \right|, x' \in B(x, h) \right\}.$$

Lemma 5 Suppose (A.1), (A.2) and (A.3) hold. Let $x \in \mathbb{R}^p$ such that $g(x) > 0$ and let $J \in \mathbb{N} \setminus \{0\}$ and $E = \{b_1, \dots, b_J\}$ with $b_1 \geq 0, \dots, b_J \geq 0$. Assume that $nh^p \bar{F}(y_n(x)|x) \rightarrow \infty$ as $n \rightarrow \infty$. If there exists $\xi > 0$ such that

$$nh^p \bar{F}(y_n(x)|x) (h \vee \Omega_n(x, h))^2 \rightarrow 0, \quad (25)$$

then, if $b_j \gamma_+(x) < 1$ for all $j = 1, \dots, J$, the random vector

$$\left\{ (nh^p \bar{F}(y_n(x)|x))^{1/2} \left(\frac{\hat{\varphi}_{b_j,n}(y_{n,j}(x)|x)}{\varphi_{b_j}(y_{n,j}(x)|x)} - 1 \right) \right\}_{j \in \{1, \dots, J\}}$$

is asymptotically Gaussian, centred, with covariance matrix $\|K\|_2^2 \Sigma_E^{(1)}(x)/g(x)$.

Proof. Let $\beta = (\beta_1, \dots, \beta_J)^t \in \mathbb{R}^J$ with $\beta \neq 0$. Our goal is to establish the asymptotic normality of the random variable

$$\Psi_n^{(1)} := \frac{1}{\Lambda_n(x)} \sum_{j=1}^J \beta_j \left(\frac{\hat{\varphi}_{b_j,n}(y_{n,j}(x)|x)}{\varphi_{b_j}(y_{n,j}(x)|x)} - 1 \right),$$

with $\Lambda_n^2(x) = 1/[nh^p \bar{F}(y_n(x)|x)]$. The following expansion holds

$$\Psi_n^{(1)} = \frac{1}{\hat{g}_n(x) \Lambda_n(x)} \sum_{j=1}^J \beta_j \left\{ g(x) \left(\frac{\hat{\psi}_{b_j,n}(y_{n,j}(x)|x)}{\psi_{b_j}(y_{n,j}(x)|x)} - 1 \right) - (\hat{g}_n(x) - g(x)) \right\}.$$

Under (A.2), it is well known that $\hat{g}_n(x) - g(x) = O(h) + O_{\mathbb{P}}((nh^p)^{-1/2})$ and thus, in particular, that $\hat{g}_n(x)$ converges in probability to $g(x)$. Hence, since $g(x) \neq 0$,

$$\frac{1}{\hat{g}_n(x) \Lambda_n(x)} (\hat{g}_n(x) - g(x)) = O\left((nh^{p+2} \bar{F}(y_n(x)|x))^{1/2}\right) + O_{\mathbb{P}}\left((\bar{F}(y_n(x)|x))^{1/2}\right) = o_{\mathbb{P}}(1).$$

As a first conclusion, the limit distribution of $\Psi_n^{(1)}$ is driven by

$$\Psi_n^{(2)} := \frac{1}{\Lambda_n(x)} \sum_{j=1}^J \beta_j \left(\frac{\hat{\psi}_{b_j,n}(y_{n,j}(x)|x)}{\psi_{b_j}(y_{n,j}(x)|x)} - 1 \right).$$

We are now interested in finding a first order expansion of $\mathbb{E}[\hat{\Psi}_{b,n}(y_{n,j}(x)|x)]$ for $b \in \mathcal{L}(E)$ with $b\gamma_+(x) < 1$ and $j = 1, \dots, J$. Since the (X_i, Y_i) , $i = 1, \dots, n$ are identically distributed,

$$\mathbb{E}[\hat{\psi}_{b,n}(y_{n,j}(x)|x)] = \int_{\mathbb{R}^p} \mathcal{K}_h(x-t) \varphi_b(y_{n,j}(x)|t) g(t) dt = \int_S \mathcal{K}(u) \varphi_b(y_{n,j}(x)|x-hu) g(x-hu) du,$$

under **(A.3)**. Hence,

$$\left| \frac{\mathbb{E}[\hat{\psi}_{b,n}(y_{n,j}(x)|x)] - \psi_b(y_{n,j}(x)|x)}{\varphi_b(y_{n,j}(x)|x)} \right| \leq \int_S \mathcal{K}(u) |g(x-hu) - g(x)| du \quad (26)$$

$$+ \int_S \mathcal{K}(u) \left| \frac{\varphi_b(y_{n,j}(x)|x-hu)}{\varphi_b(y_{n,j}(x)|x)} - 1 \right| g(x-hu) du. \quad (27)$$

Under **(A.2)**, we have

$$(26) \leq c_g h \int_S d(u, 0) \mathcal{K}(u) du = O(h). \quad (28)$$

Besides, in view of (28),

$$(27) \leq \Omega_n(x, h) \int_S \mathcal{K}(u) g(x-hu) du \leq g(x) \Omega_n(x, h) (1 + o(1)). \quad (29)$$

Combining (28) and (29) leads to

$$\frac{\mathbb{E}[\hat{\psi}_{b,n}(y_{n,j}(x)|x)]}{\psi_b(y_{n,j}(x)|x)} - 1 = O(h) + O(\Omega_n(x, h)) \quad (30)$$

for all $b \in \mathcal{L}(E)$ and $j = 1, \dots, J$. Thus, by assumption

$$\frac{1}{\Lambda_n(x)} \left(\frac{\mathbb{E}[\hat{\psi}_{b_j,n}(y_{n,j}(x)|x)]}{\psi_{b_j}(y_{n,j}(x)|x)} - 1 \right) \rightarrow 0,$$

as $n \rightarrow \infty$ and for all $j = 1, \dots, J$. The limiting distribution of $\Psi_n^{(2)}$ (and consequently the one of $\Psi_n^{(1)}$) is then also the limiting distribution of

$$\Psi_n := \sum_{j=1}^J \beta_j \left(\frac{\hat{\psi}_{b_j,n}(y_{n,j}(x)|x) - \mathbb{E}[\hat{\psi}_{b_j,n}(y_{n,j}(x)|x)]}{\Lambda_n(x) \psi_{b_j}(y_{n,j}(x)|x)} \right) =: \sum_{i=1}^n Z_{i,n},$$

where

$$Z_{i,n} := \frac{1}{n \Lambda_n(x)} \sum_{j=1}^J \beta_j \frac{\left\{ \mathcal{K}_h(x - X_i) Y_i^{b_j} \mathbb{I}\{Y_i \geq y_{n,j}(x)\} - \mathbb{E}(\mathcal{K}_h(x - X_i) Y^{b_j} \mathbb{I}\{Y \geq y_{n,j}(x)\}) \right\}}{\psi_{b_j}(y_{n,j}(x)|x)}.$$

Clearly, $\{Z_{i,n}, i = 1, \dots, n\}$ is a set of centred, independent and identically distributed random variables with variance $\mathbb{V}(Z_{1,n}) = \beta^t B \beta / (n \Lambda_n(x))^2$ where B is the $J \times J$ matrix defined for $(j, l) \in \{1, \dots, J\}^2$ by

$$B_{j,l} := \text{cov} \left(\mathcal{K}_h(x - X) \frac{Y^{b_j} \mathbb{I}\{Y \geq y_{n,j}(x)\}}{\psi_{b_j}(y_{n,j}(x)|x)}, \mathcal{K}_h(x - X) \frac{Y^{b_l} \mathbb{I}\{Y \geq y_{n,l}(x)\}}{\psi_{b_l}(y_{n,l}(x)|x)} \right).$$

Introducing the function $\mathcal{Q}(\cdot) := \mathcal{K}^2(\cdot) / \|\mathcal{K}\|_2^2$, one has

$$\begin{aligned} B_{j,l} &= \frac{1}{h^p \psi_{b_j}(y_{n,j}(x)|x) \psi_{b_l}(y_{n,l}(x)|x)} \left\{ \|\mathcal{K}\|_2^2 \mathbb{E}(\mathcal{Q}_h(x - X) Y^{b_j+b_l} \mathbb{I}\{Y \geq y_{n,j}(x) \vee y_{n,l}(x)\}) \right. \\ &\quad \left. - h^p \mathbb{E}(\mathcal{K}_h(x - X) Y^{b_j} \mathbb{I}\{Y \geq y_{n,j}(x)\}) \mathbb{E}(\mathcal{K}_h(x - X) Y^{b_l} \mathbb{I}\{Y \geq y_{n,l}(x)\}) \right\}. \end{aligned}$$

Since $\mathcal{Q}(\cdot)$ also satisfies assumption **(A.3)**, the three above expectations can be expanded as in (30) leading to:

$$B_{j,l} = \frac{\|\mathcal{K}\|_2^2 \psi_{b_j+b_l}(y_{n,j}(x) \vee y_{n,l}(x)|x)}{h^p \psi_{b_j}(y_{n,j}(x)|x) \psi_{b_l}(y_{n,l}(x)|x)} \left[1 + O \left(h^p \frac{\psi_{b_j}(y_{n,j}(x)|x) \psi_{b_l}(y_{n,l}(x)|x)}{\psi_{b_j+b_l}(y_{n,j}(x) \vee y_{n,l}(x)|x)} \right) \right].$$

Recall that $\psi_b(\cdot|x) = g(x)\varphi_b(\cdot|x)$. Since $(b_j + b_l)\gamma_+(x) < 1$ and $y_{n,j}(x) = y_n(x)(1 + o(1))$ for all $j = 1, \dots, J$, Lemma 2(i) entails that

$$\begin{aligned} \frac{\psi_{b_j}(y_{n,j}(x)|x)\psi_{b_l}(y_{n,l}(x)|x)}{\psi_{b_j+b_l}(y_{n,j}(x) \vee y_{n,l}(x)|x)} &= \frac{1 - (b_j + b_l)\gamma_+(x)}{(1 - b_j\gamma_+(x))(1 - b_l\gamma_+(x))} \frac{\bar{F}(y_{n,j}|x)\bar{F}(y_{n,l}|x)}{\bar{F}(y_{n,j}(x) \vee y_{n,l}(x)|x)}(1 + o(1)) \\ &= \frac{1 - (b_j + b_l)\gamma_+(x)}{(1 - b_j\gamma_+(x))(1 - b_l\gamma_+(x))} \bar{F}(y_{n,j}(x) \wedge y_{n,l}(x)|x)(1 + o(1)), \end{aligned}$$

and consequently,

$$B_{j,l} = \frac{\|\mathcal{K}\|_2^2}{h^p g(x)} \frac{(1 - b_j\gamma_+(x))(1 - b_l\gamma_+(x))}{1 - (b_j + b_l)\gamma_+(x)} \frac{1}{\bar{F}(y_{n,j}(x) \wedge y_{n,l}(x)|x)}(1 + o(1)).$$

Moreover, since $\bar{F}(y_{n,j}(x) \wedge y_{n,l}(x)|x) = \bar{F}(y_n(x) + (t_{n,j}(x) \vee t_{n,l}(x))a(y_n(x)|x))$, Lemma 1, equation (18) entails that

$$\lim_{n \rightarrow \infty} \frac{\bar{F}(y_{n,j}(x) \wedge y_{n,l}(x)|x)}{\bar{F}(y_n(x)|x)} \xrightarrow{n \rightarrow \infty} \frac{1}{L_{\gamma(x)}^{\leftarrow}(0)} = 1, \quad (31)$$

leading to $B_{j,l} = n\Lambda_n^2(x)\|\mathcal{K}\|_2^2[\Sigma_E^{(1)}(x)]_{j,l}/g(x)(1 + o(1))$.

To sum up, the variance of Ψ_n converges to $\|\mathcal{K}\|_2^2\beta^t\Sigma_E^{(1)}(x)\beta/g(x)$. Since Ψ_n is a sum of independent centred random values, its asymptotic normality can be established using Lyapounov theorem. It is sufficient to prove that there exists $\eta > 0$ such that:

$$\sum_{i=1}^n \mathbb{E} \left(|Z_{i,n}|^{2+\eta} \right) = n\mathbb{E} \left(|Z_{1,n}|^{2+\eta} \right) \rightarrow 0.$$

Since for every random pair (T_1, T_2) with finite $(2+\eta)$ th order moments, one has $\mathbb{E}(|T_1 + T_2|^{2+\eta}) \leq 2^{2+\eta}[\mathbb{E}(|T_1|^{2+\eta}) \vee \mathbb{E}(|T_2|^{2+\eta})]$,

$$\begin{aligned} \mathbb{E}|Z_{1,n}|^{2+\eta} &= \left(\frac{1}{n\Lambda_n(x)} \right)^{2+\eta} \mathbb{E} \left[\left| \sum_{j=1}^J \frac{\beta_j \mathcal{K}_h(x - X) Y^{b_j} \mathbb{I}\{Y \geq y_{n,j}(x)\}}{\psi_{b_j}(y_{n,j}(x)|x)} \right|^{2+\eta} \right] \\ &= \mathbb{E} \left(\left| \sum_{j=1}^J \frac{\beta_j \mathcal{K}_h(x - X) Y^{b_j} \mathbb{I}\{Y \geq y_{n,j}(x)\}}{\psi_{b_j}(y_{n,j}(x)|x)} \right|^{2+\eta} \right) \\ &\leq \left(\frac{2}{n\Lambda_n(x)} \right)^{2+\eta} \mathbb{E} \left[\left(\sum_{j=1}^J \frac{|\beta_j| \mathcal{K}_h(x - X) Y^{b_j} \mathbb{I}\{Y \geq y_{n,j}(x)\}}{\psi_{b_j}(y_{n,j}(x)|x)} \right)^{2+\eta} \right]. \end{aligned}$$

From Lemma 2(i) and (31), $\psi_{b_j}(y_{n,j}(x)|x) = g(x)y_n^{b_j}(x)\bar{F}(y_n(x)|x)/(1 - b_j\gamma_+(x))(1 + o(1))$ which implies that for n large enough,

$$\mathbb{E}|Z_{1,n}|^{2+\eta} \leq \left(\frac{4h^p\Lambda_n(x)}{g(x)} \right)^{2+\eta} \mathbb{E} \left[\left(\sum_{j=1}^J |\beta_j| \mathcal{K}_h(x - X) \mathbb{I}\{Y \geq y_{n,j}(x)\} \left(\frac{Y}{y_n(x)} \right)^{b_j} (1 - b_j\gamma_+(x)) \right)^{2+\eta} \right].$$

Letting $\check{b} = \min\{b_1, \dots, b_J\}$, $\tilde{b} = \max\{b_1, \dots, b_J\}$ and $\check{y}_n(x) = \min\{y_{n,1}(x), \dots, y_{n,J}(x)\}$, it follows that

$$\mathbb{E}|Z_{1,n}|^{2+\eta} \leq \frac{4^{2+\eta}}{g(x)} \left(\frac{\Lambda_n(x)(1 - \check{b}\gamma_+(x))}{y_n^{\check{b}}(x)} \right)^{2+\eta} \left(\sum_{j=1}^J |\beta_j| \right)^{2+\eta} \mathbb{E} \left[\left(\mathcal{K} \left(\frac{x - X}{h} \right) Y^{\tilde{b}} \mathbb{I}\{Y \geq \check{y}_n(x)\} \right)^{2+\eta} \right].$$

Let $\mathcal{N}(\cdot) := \mathcal{K}^{2+\eta}(\cdot)/\|\mathcal{K}\|_{2+\eta}^{2+\eta}$ and choose η such that $0 < \eta < -2 + 1/(\tilde{b}\gamma_+(x))$. One has

$$\begin{aligned} \mathbb{E} \left[\left(\mathcal{K} \left(\frac{x-X}{h} \right) Y^{\tilde{b}} \mathbb{I}\{Y \geq \check{y}_n(x)\} \right)^{2+\eta} \right] &= h^p \|\mathcal{K}\|_{2+\eta}^{2+\eta} \mathbb{E} \left(\mathcal{N}_h(x-X) Y^{\tilde{b}(2+\eta)} \mathbb{I}\{Y \geq \check{y}_n(x)\} \right), \\ &= h^p \|\mathcal{K}\|_{2+\eta}^{2+\eta} \psi_{\tilde{b}(2+\eta)}(\check{y}_n(x)|x)(1+o(1)), \end{aligned}$$

using expansion (30) that holds since $\mathcal{N}(\cdot)$ also fulfills assumption **(A.3)**. Lemma 2(i) and (31) entail $n\mathbb{E}|Z_{1,n}|^{2+\eta} = O(\Lambda_n^\eta(x)) \rightarrow 0$ as $n \rightarrow \infty$ which concludes the proof. \blacksquare

The next result establishes the asymptotic behavior of the random vector $\{\hat{\varphi}_{b_j,n}^\leftarrow(\alpha_{n,j}|x)\}_{j=1,\dots,J}$ with for all $j = 1, \dots, J$, $\alpha_{n,j} = \varphi_{b_j}[\text{RVaR}(\alpha_n|x)|x](1+o(1))$ where $\alpha_n \rightarrow 0$ as $n \rightarrow \infty$.

Lemma 6 *Suppose **(A.1)**, **(A.2)** and **(A.3)** hold. Let $x \in \mathbb{R}^p$ such that $g(x) > 0$ and let $J \in \mathbb{N} \setminus \{0\}$, $E = \{b_1, \dots, b_J\}$ with $b_1 \geq 0, \dots, b_J \geq 0$. Assume $\alpha_n \rightarrow 0$ and $nh^p\alpha_n \rightarrow \infty$ as $n \rightarrow \infty$. If there exists $\xi > 0$ such that*

$$nh^p\alpha_n \left(h \vee \max_{b \in \mathcal{L}(E)} \omega(x, \alpha_n, b, \xi, h) \right)^2 \rightarrow 0,$$

then, if $2b_j\gamma_+(x) < 1$, for all $j = 1, \dots, J$, the random vector

$$\left\{ \frac{\text{RVaR}(\alpha_n|x)}{a(\text{RVaR}(\alpha_n|x)|x)} (nh^p\alpha_n)^{1/2} \left(\frac{\hat{\varphi}_{b_j,n}^\leftarrow(\alpha_{n,j}|x)}{\varphi_{b_j}^\leftarrow(\alpha_{n,j}|x)} - 1 \right) \right\}_{j \in \{1, \dots, J\}}$$

is asymptotically Gaussian, centred, with covariance matrix $\|\mathcal{K}\|_2^2 \Sigma_E^{(3)}(x)/g(x)$ where $[\Sigma_E^{(3)}(x)]_{i,j} = [1 - (b_i + b_j)\gamma_+(x)]^{-1}$, for all $(i, j) \in \{1, \dots, J\}^2$.

Proof. Let $(z_1, \dots, z_J) \in \mathbb{R}^J$. We are interested in the asymptotic behavior of the c.d.f. defined by

$$\Phi_n(z_1, \dots, z_J) = \mathbb{P} \left(\bigcap_{j=1}^J \left\{ \sigma_{n,j}^{-1}(x) (\hat{\varphi}_{b_j,n}^\leftarrow(\alpha_{n,j}|x) - \varphi_{b_j}^\leftarrow(\alpha_{n,j}|x)) \leq z_j \right\} \right),$$

where for $j = 1, \dots, J$, $\sigma_{n,j}^{-1}(x) := (1 - b_j\gamma_+(x))(nh^p\alpha_n)^{1/2}/a(\text{RVaR}(\alpha_n|x)|x)$. For $j = 1, \dots, J$, let us introduce the notations

$$W_{n,j}(x) = \frac{(nh^p\alpha_n)^{1/2}}{\alpha_{n,j}} \left(\hat{\varphi}_{b_j,n}(\varphi_{b_j}^\leftarrow(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j)|x) - \varphi_{b_j}(\varphi_{b_j}^\leftarrow(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j)|x) \right)$$

and

$$s_{n,j}(x) = \frac{(nh^p\alpha_n)^{1/2}}{\alpha_{n,j}} \left(\alpha_{n,j} - \varphi_{b_j}(\varphi_{b_j}^\leftarrow(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j)|x) \right).$$

It is easy to check that

$$\Phi_n(z_1, \dots, z_J) = \mathbb{P} \left(\bigcap_{j=1}^J \{W_{n,j}(x) \leq s_{n,j}(x)\} \right).$$

Let us first focus on the non-random term $s_{n,j}(x)$ for $j \in \{1, \dots, J\}$. Since, for all $b \geq 0$ such that $b\gamma_+(x) < 1$, the function $\varphi_b(\cdot|x)$ is continuously differentiable, there exists $\theta_{n,j} \in (0, 1)$ such that

$$\varphi_{b_j} \left(\varphi_{b_j}^\leftarrow(\alpha_{n,j}|x)|x) \right) - \varphi_{b_j} \left(\varphi_{b_j}^\leftarrow(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j|x) \right) = -\sigma_{n,j}(x)z_j\varphi'_{b_j}(\varphi_{b_j}^\leftarrow(\alpha_{n,j}|x)|x), \quad (32)$$

where $r_{n,j}(x) := \varphi_{b_j}^{\leftarrow}(\alpha_{n,j}|x) + \theta_{n,j}\sigma_{n,j}(x)z_j = \text{RVaR}(\alpha_n|x) + \tilde{t}_{n,j}(x; z_j)a(\text{RVaR}(\alpha_n|x)|x)$ with

$$\tilde{t}_{n,j}(x; \theta_{n,j}) := \frac{\varphi_{b_j}^{\leftarrow}(\alpha_{n,j}|x) - \text{RVaR}(\alpha_n|x)}{a(\text{RVaR}(\alpha_n|x)|x)} + \frac{\theta_{n,j}z_j}{1 - b\gamma_+(x)}(nh^p\alpha_n)^{-1/2}.$$

Our goal is now to find a first order expansion of $\varphi_{b_j}'(r_{n,j}(x)|x)$. Since from Lemma 4, $\varphi_b^{\leftarrow}(\cdot|x)$ satisfies the assumption of Lemma 1 for all $b \geq 0$ such that $b\gamma_+(x) < 1$, equation (17) of Lemma 1 entails that

$$\frac{\varphi_{b_j}^{\leftarrow}(\alpha_{n,j}|x) - \text{RVaR}(\alpha_n|x)}{a(\text{RVaR}(\alpha_n|x)|x)} \rightarrow 0. \quad (33)$$

Using (33), one has for all $j \in \{1, \dots, J\}$ that $\tilde{t}_{n,j}(x; z_j) \rightarrow 0$. Hence, from equation (18) of Lemma 1,

$$\frac{\varphi_{b_j}(r_{n,j}(x)|x)}{\varphi_{b_j}(\text{RVaR}(\alpha_n|x)|x)} \rightarrow 1. \quad (34)$$

Since $a(y|x)/y \rightarrow \gamma_+(x)$ as $y \uparrow y^*(x)$, it is thus clear that $r_{n,j}(x) = \text{RVaR}(\alpha_n|x)(1 + o(1)) \uparrow y^*(x)$. Using Lemma 2(ii) and (34), it follows that

$$\varphi_{b_j}'(r_{n,j}(x)|x) = (b_j\gamma_+(x) - 1) \frac{\varphi_{b_j}(\text{RVaR}(\alpha_n|x)|x)}{a(r_{n,j}(x)|x)}(1 + o(1)). \quad (35)$$

Let us focus on the sequence $a(r_{n,j}(x)|x)$. Under **(A.1)**, using equation (18) of Lemma 1, it is clear that $\bar{F}(r_{n,j}(x)|x) = \alpha_n(1 + o(1))$ for all $j = 1, \dots, J$. In addition, the function $f(\text{RVaR}(\cdot|x)|x)$ is regularly varying and thus, $f(r_{n,j}(x)|x) = f(\text{RVaR}(\alpha_n|x)|x)(1 + o(1))$. Hence,

$$a(r_{n,j}(x)|x) = \frac{\bar{F}(r_{n,j}(x)|x)}{f(r_{n,j}(x)|x)} = \frac{\alpha_n}{f(\text{RVaR}(\alpha_n|x)|x)}(1 + o(1)) = a(\text{RVaR}(\alpha_n|x)|x)(1 + o(1)).$$

Substituting in (35) the sequence $a(r_{n,j}(x)|x)$ by its first order approximation, we obtain that

$$\varphi_{b_j}'(r_{n,j}(x)|x) = (b_j\gamma_+(x) - 1) \frac{\varphi_{b_j}(\text{RVaR}(\alpha_n|x)|x)}{a(\text{RVaR}(\alpha_n|x)|x)}(1 + o(1)). \quad (36)$$

To conclude the study of the non-random term $s_{n,j}(x)$, it suffices to replace in (32), $\varphi_{b_j}'(r_{n,j}(x)|x)$ by its first order approximation, leading to

$$s_{n,j}(x) = -z_j \frac{\sigma_{n,j}(x)(b_j\gamma_+(x) - 1)}{a(\text{RVaR}(\alpha_n|x)|x)(nh^p\alpha_n)^{-1/2}} \frac{\varphi_{b_j}(\text{RVaR}(\alpha_n|x)|x)}{\alpha_{n,j}}(1 + o(1)) = z_j(1 + o(1)). \quad (37)$$

Let us now turn to the random term $W_{n,j}(x)$ for $j \in \{1, \dots, J\}$. Let $y_n(x) := \text{RVaR}(\alpha_n|x)$ and for all $j = 1, \dots, J$, $y_{n,j}(x) := \varphi_{b_j}^{\leftarrow}(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j = \text{RVaR}(\alpha_n|x) + \tilde{t}_{n,j}(x; 1)a(\text{RVaR}(\alpha_n|x)|x)$. Obviously, equation (34) also holds if $r_{n,j}(x)$ is replaced by $y_{n,j}(x)$ and thus,

$$\varphi_{b_j}(y_{n,j}(x)|x) = \alpha_{n,j}(1 + o(1)) = \varphi_{b_j}(\text{RVaR}(\alpha_n|x)|x)(1 + o(1)).$$

As a consequence, there exists $\xi \in (0, 1)$ such that for n large enough

$$\Omega_n(x, h) \leq \max_{b \in \mathcal{L}(E)} \omega(x, \alpha_n, b, \xi, h).$$

One can then apply Lemma 5 to show that the random vector $(W_{n,1}, \dots, W_{n,J})$ is equal to ξ_n where ξ_n is a J -random vector converging to a centred Gaussian random variable with covariance

matrix $\|\mathcal{K}\|_2^2 \Sigma_E^{(1)}(x)/g(x)$. Taking account of (37) and since, as a consequence of (33), $\varphi_{b_j}^{\leftarrow}(\alpha_{n,j}|x) = \text{RVaR}(\alpha_n|x)(1+o(1))$, we have shown that

$$\left\{ \frac{\text{RVaR}(\alpha_n|x)(nh^p\alpha_n)^{1/2}}{a(\text{RVaR}(\alpha_n|x)|x)} \left(\frac{\hat{\varphi}_{b_j,n}^{\leftarrow}(\alpha_{n,j}|x)}{\varphi_{b_j}^{\leftarrow}(\alpha_{n,j}|x)} - 1 \right) \right\}_{j=1,\dots,J}$$

converges to the c.d.f. of a centred Gaussian distribution with covariance matrix $\|\mathcal{K}\|_2^2 \Sigma_E^{(3)}(x)/g(x) = \|\mathcal{K}\|_2^2 A(x) \Sigma_E^{(1)}(x) A(x)/g(x)$ where $A(x)$ is a diagonal matrix with $A_{j,j}(x) = (1 - b_j \gamma_+(x))^{-1}$. ■

The next result will be used in the proofs of Corollary 2 and 3. The following notations are introduced. For $x \in \mathbb{R}^J$, the gradient of the function $\Psi : \mathbb{R}^J \mapsto \mathbb{R}$ evaluated at x is denoted by $(\nabla \Psi)_x$ and the $J \times J$ matrix D_x is the diagonal matrix whose (diagonal) elements are the coordinates of x .

Lemma 7 *Let $J \in \mathbb{N} \setminus \{0\}$ and $\Psi : \mathbb{R}^J \mapsto \mathbb{R}$ be a continuously differentiable function. Assume that the assumptions of Theorem 2 hold with $E = \{b_1, \dots, b_J\}$ where $b_1 \geq 0, \dots, b_J \geq 0$. Introduce $r_{n,E} := \{\text{RCTM}_{b_i}(\alpha_n|x)\}_{i \in \{1, \dots, J\}}$ and $\hat{r}_{n,E} := \{\widehat{\text{RCTM}}_{b_i,n}(\alpha_n|x)\}_{i \in \{1, \dots, J\}}$. If there exist a positive sequence v_n and a non null vector $v \in \mathbb{R}^J$ such that,*

$$v_n D_{r_{n,E}} (\nabla \Psi)_{\theta \hat{r}_{n,E} + (1-\theta)r_{n,E}} \xrightarrow{P} v,$$

for all $\theta \in (0, 1)$, then

$$v_n (n \ell^p \alpha_n)^{1/2} [\Psi(\hat{r}_{n,E}) - \Psi(r_{n,E})]$$

is asymptotically Gaussian, centred, with variance $\|\mathcal{K}\|_2^2 (v^t \Sigma_E^{(1)}(x) v)/g(x)$ in the case $h/k \rightarrow 0$ and $\|\mathcal{K}\|_2^2 (v^t \Sigma_E^{(2)}(x) v)/g(x)$ in the case $k/h \rightarrow 0$.

Proof. A first order Taylor expansion leads to

$$\Psi(\hat{r}_{n,E}) - \Psi(r_{n,E}) = (\hat{r}_{n,E} - r_{n,E})^t (\nabla \Psi)_{\theta \hat{r}_{n,E} + (1-\theta)r_{n,E}},$$

where $\theta \in (0, 1)$. From Theorem 2, one has $\hat{r}_{n,E} - r_{n,E} = (n \ell^p \alpha_n)^{-1/2} D_{r_{n,E}} \xi_n$, where ξ_n is a random vector of dimension J asymptotically Gaussian, centred, with covariance matrix $\|\mathcal{K}\|_2^2 \Sigma_E^{(1)}(x)/g(x)$ if $h/k \rightarrow 0$ and $\|\mathcal{K}\|_2^2 \Sigma_E^{(2)}(x)/g(x)$ if $k/h \rightarrow 0$. Hence, by assumption

$$v_n (n \ell^p \alpha_n)^{1/2} (\Psi(\hat{r}_{n,E}) - \Psi(r_{n,E})) = v_n \xi_n^t D_{r_{n,E}} (\nabla \Psi)_{\theta \hat{r}_{n,E} + (1-\theta)r_{n,E}} = v^t \xi (1 + o_{\mathbb{P}}(1)),$$

and the proof is complete. ■

7.2 Proofs of main results

Proof of Proposition 1. (i) Conditionally to $\{X = x\}$, $Y \leq y^*(x)$ a.s. and thus

$$\frac{\text{RCTM}_b(\alpha|x)}{[y^*(x)]^b} = \frac{1}{\alpha} \mathbb{E} \left[\left(\frac{Y}{y^*(x)} \right)^b \mathbb{I}\{Y > \text{RVaR}(\alpha|x)\} \middle| X = x \right] \leq 1.$$

Moreover,

$$\frac{\text{RCTM}_b(\alpha|x)}{[y^*(x)]^b} \geq \frac{1}{\alpha} \left(\frac{\text{RVaR}(\alpha|x)}{y^*(x)} \right)^b \mathbb{P}(Y > \text{RVaR}(\alpha|x) | X = x) = \left(\frac{\text{RVaR}(\alpha|x)}{y^*(x)} \right)^b \rightarrow 1,$$

since $\text{RVaR}(\alpha|x) \rightarrow y^*(x)$ as $\alpha \rightarrow 0$. These two bounds prove the result.

(ii) Recall that $\text{RVaR}(\alpha|x) = \varphi_0^{\leftarrow}(\alpha|x)$ and $\text{RCTM}_b(\alpha|x) = \varphi_b(\varphi_0^{\leftarrow}(\alpha|x)|x)/\alpha$. Then, applying Lemma 2(i) with $y = \varphi_0^{\leftarrow}(\alpha|x)$ entails that

$$\lim_{\alpha \rightarrow 0} \frac{\varphi_b(\varphi_0^{\leftarrow}(\alpha|x)|x)}{\alpha[\varphi_0^{\leftarrow}(\alpha|x)]^b} = \frac{1}{1 - b\gamma_+(x)},$$

since $\bar{F}(\cdot|x) = \varphi_0(\cdot|x)$, and (9) is proved. Next, Lemma 1, (17) with $U(\cdot|x) = \text{RVaR}(\cdot|x)$ implies

$$\lim_{\alpha \rightarrow 0} \frac{\text{RVaR}(s\alpha|x)/\text{RVaR}(\alpha|x) - 1}{a(\text{RVaR}(\alpha|x)|x)/\text{RVaR}(\alpha|x)} = L_{\gamma(x)}(1/s). \quad (38)$$

Two cases occur. If $\gamma(x) > 0$, (38) and (19) entail that

$$\lim_{\alpha \rightarrow 0} \frac{\text{RVaR}(s\alpha|x)}{\text{RVaR}(\alpha|x)} = 1 + \gamma(x)L_{\gamma(x)}(1/s) = s^{-\gamma(x)}.$$

If $\gamma(x) \leq 0$, (38) and (19) yield

$$\lim_{\alpha \rightarrow 0} \frac{\text{RVaR}(s\alpha|x)}{\text{RVaR}(\alpha|x)} = 1,$$

and the proof is complete. ■

Proof of Theorem 1. Remarking that $a(\cdot|x) = \bar{F}(\cdot|x)/f(\cdot|x)$ (see Table 1), the result is a direct consequence of Lemma 6 with $J = 1$ and $b_1 = 0$. ■

Proof of Theorem 2. Let $(z_1, \dots, z_J) \in \mathbb{R}^J$. Our goal is to prove that the c.d.f. defined by

$$\Phi_n(z_1, \dots, z_J) = \mathbb{P} \left(\bigcap_{j=1}^J \left\{ \sigma_{n,j}^{-1}(x) (\widehat{\text{RCTM}}_{b_j,n}(\alpha_n|x) - \text{RCTM}_{b_j}(\alpha_n|x) \leq z_j \right\} \right),$$

with $\sigma_{n,j}(x) := \text{RCTM}_{b_j}(\alpha_n|x)(n\ell^p\alpha_n)^{-1/2}$ converges to the c.d.f. of a Gaussian random vector. To this aim, introduce the following notations: for $j = 1, \dots, J$ and $\theta > 0$ let $\alpha_{n,j}(\theta) := \alpha_n(\text{RCTM}_{b_j}(\alpha_n|x) + \theta\sigma_{n,j}(x)z_j)$, $v_{n,j}(x) := (1 - b_j\gamma_+(x))(n\ell^p\alpha_n)^{1/2}/a(\text{RVaR}(\alpha_n|x)|x)$,

$$W_{n,j}(x) = v_{n,j}(x) \left(\hat{\varphi}_{b_j,n}^{\leftarrow}(\alpha_{n,j}(1)|x) - \varphi_{b_j}^{\leftarrow}(\alpha_{n,j}(1)|x) \right),$$

and

$$W_{n,j}^{(0)}(x) = v_{n,j}(x) \left(\widehat{\text{RVaR}}_n(\alpha_n|x) - \text{RVaR}(\alpha_n|x) \right).$$

It is easy to check that

$$\Phi_n(z_1, \dots, z_J) = \mathbb{P} \left(\bigcap_{j=1}^J \left\{ W_{n,j}(x) - W_{n,j}^{(0)}(x) \leq s_{n,j}(x) \right\} \right), \quad (39)$$

with $s_{n,j}(x) := v_{n,j}(x)(\text{RVaR}(\alpha_n|x) - \varphi_{b_j}^{\leftarrow}(\alpha_{n,j}(1)|x))$. Let us first focus on the non-random term $s_{n,j}(x)$ for $j = \{1, \dots, J\}$. Remarking that $\text{RVaR}(\alpha_n|x) = \varphi_{b_j}^{\leftarrow}(\alpha_n \text{RCTM}_{b_j}(\alpha_n|x)|x)$, a first order Taylor expansion leads to: $s_{n,j}(x) = -z_j\alpha_nv_{n,j}(x)\sigma_{n,j}(x)(\varphi_{b_j}^{\leftarrow})'(\alpha_{n,j}(\theta_{n,j})|x)$ where

$(\theta_{n,1}, \dots, \theta_{n,J}) \in (0, 1)^J$.

Our aim is to find a first order expansion of $(\varphi_{b_j}^{\leftarrow})'(\alpha_{n,j}(\theta_{n,j})|x) = 1/(\varphi_{b_j}'(\varphi_{b_j}^{\leftarrow}(\alpha_{n,j}(\theta_{n,j})|x)|x)$. First note that $\varphi_{b_j}^{\leftarrow}(\alpha_{n,j}(\theta_{n,j})|x) = \text{RVaR}(\alpha_n|x) + t_{n,j}(x)a(\text{RVaR}(\alpha_n|x)|x)$ with

$$t_{n,j}(x) = \frac{\varphi_{b_j}^{\leftarrow}(\alpha_{n,j}(\theta_{n,j})|x) - \text{RVaR}(\alpha_n|x)}{a(\text{RVaR}(\alpha_n|x)|x)}.$$

Since $\alpha_{n,j}(\theta_{n,j}) = \varphi_{b_j}(\text{RVaR}(\alpha_n|x)|x)(1+o(1))$, equation (33) in the proof of Lemma 6 entails that $t_{n,j}(x) \rightarrow 0$. Hence, equation (36) also holds if $r_{n,j}(x)$ is replaced by $\varphi_{b_j}^{\leftarrow}(\alpha_{n,j}(\theta_{n,j})|x)$. Therefore,

$$(\varphi_{b_j}^{\leftarrow})'(\alpha_{n,j}(\theta_{n,j})|x) = \frac{a(\text{RVaR}(\alpha_n|x)|x)}{(b_j\gamma_+(x) - 1)\varphi_{b_j}(\text{RVaR}(\alpha_n|x)|x)}(1 + o(1)).$$

Replacing in the Taylor expansion of $s_{n,j}(x)$ yields $s_{n,j}(x) = z_j(1 + o(1))$.

We are now interested in the random term

$$W_{n,j}(x) = \frac{1 - b_j\gamma_+(x)}{a(\text{RVaR}(\alpha_n|x)|x)}(n\ell^p\alpha_n)^{1/2} \left(\hat{\varphi}_{b_j,n}^{\leftarrow}(\alpha_{n,j}(1)|x) - \varphi_{b_j}^{\leftarrow}(\alpha_{n,j}(1)|x) \right).$$

Since $\varphi_{b_j}^{\leftarrow}(\alpha_{n,j}(1)|x) = \text{RVaR}(\alpha_n|x)(1+o(1))$ and $\alpha_{n,j}(1) = \varphi_{b_j}(\text{RVaR}(\alpha_n|x)|x)(1+o(1))$, Lemma 6 entails that

$$\{W_{n,j}\}_{j=1,\dots,J} = (\ell/h)^{p/2} \xi_n, \quad (40)$$

where ξ_n is asymptotically Gaussian, centred with covariance matrix $\|\mathcal{K}\|_2^2 \tilde{A}(x) \Sigma_E^{(3)}(x) \tilde{A}(x)/g(x) = \|\mathcal{K}\|_2^2 \Sigma_E^{(1)}(x)/g(x)$ and where $\tilde{A}(x)$ is a diagonal matrix with $\tilde{A}_{j,j}(x) = 1 - b_j\gamma_+(x)$. Furthermore, from Theorem 1,

$$\left\{ W_{n,j}^{(0)} \right\}_{j=1,\dots,J} = (\ell/k)^{p/2} \xi_n^{(0)}, \quad (41)$$

where $\xi_n^{(0)}$ is asymptotically Gaussian, centred with covariance $\|\mathcal{K}\|_2^2 v v^t / g(x) = \|\mathcal{K}\|_2^2 \Sigma_E^{(2)}(x)/g(x)$ and $v = (1 - b_1\gamma_+(x), \dots, 1 - b_J\gamma_+(x))^t \in \mathbb{R}^J$. Collecting (39)–(41) and using the fact that $s_{n,j}(x) \rightarrow z_j$ conclude the proof. \blacksquare

Proof of Corollary 2. First, recall the notations introduced in Lemma 7: For all set $E = \{b_1, \dots, b_J\}$ where $b_1 \geq 0, \dots, b_J \geq 0$, $r_{n,E} = (\text{RCTM}_{b_1}(\alpha_n|x), \dots, \text{RCTM}_{b_J}(\alpha_n|x))^t$ and $\hat{r}_{n,E} = (\widehat{\text{RCTM}}_{b_1,n}(\alpha_n|x), \dots, \widehat{\text{RCTM}}_{b_J,n}(\alpha_n|x))^t$. Remark also that under the assumptions of Theorem 2 with $E = \{b_1, \dots, b_J\}$, one has for all $\theta \in (0, 1)$ and $i \in \{1, \dots, J\}$ that

$$\theta \widehat{\text{RCTM}}_{i,n}(\alpha_n|x) + (1 - \theta) \text{RCTM}_i(\alpha_n|x) = \text{RCTM}_i(\alpha_n|x) (1 + O_{\mathbb{P}}(\sigma_n)), \quad (42)$$

where $\sigma_n := (n\ell^p\alpha_n)^{-1/2} \rightarrow 0$.

(i) is a direct consequence of Theorem 2 with $E = \{1\}$.

(ii) Let $E = \{1, 2\}$. We start by remarking that $\text{RCTV}(\alpha_n|x) = \Psi(r_{n,E})$ where for $x = (x_1, x_2)^t$, $\Psi(x) = x_2 - x_1^2$. Since $(\nabla \Psi)_x = (-2x_1, 1)^t$, equation (42) entails that

$$D_{r_{n,E}}(\nabla \Psi)_{\theta \hat{r}_{n,E} + (1-\theta)r_{n,E}} = (-2[\text{RCTM}_1(\alpha_n|x)]^2, \text{RCTM}_2(\alpha_n|x))^t (1 + O_{\mathbb{P}}(\sigma_n)).$$

Proposition 1(ii) leads to $D_{r_{n,E}}(\nabla \Psi)_{\theta \hat{r}_{n,E} + (1-\theta)r_{n,E}} = [\text{RVaR}(\alpha_n|x)]^2 v^t (1 + o_{\mathbb{P}}(1))$ with

$$v = \left(-\frac{2}{(1 - \gamma_+(x))^2}, \frac{1}{1 - 2\gamma_+(x)} \right)^t.$$

Thus, the assumptions of Lemma 7 are satisfied with $v_n = [\text{RVaR}(\alpha_n|x)]^{-2}$. The end of the proof is straightforward by using Lemma 7 and remarking that

$$\begin{aligned} v^t \Sigma_E^{(1)}(x)v &= \frac{1 - 7\gamma_+(x) + 9\gamma_+^2(x) + 15\gamma_+^3(x) - 6\gamma_+^4(x)}{(1 - \gamma_+(x))^2(1 - 2\gamma_+(x))(1 - 3\gamma_+(x))(1 - 4\gamma_+(x))} = \vartheta_{\text{RCTV},1}(\gamma_+(x)) \\ \text{and } v^t \Sigma_E^{(2)}(x)v &= \left(\frac{1 + \gamma_+(x)}{1 - \gamma_+(x)} \right)^2 = \vartheta_{\text{RCTV},2}(\gamma_+(x)). \end{aligned}$$

(iii) Let $E = \{1, 2, 3\}$ and remark that $\text{RCTS}(\alpha_n|x) = \Psi(r_{n,E})$ where for $x = (x_1, x_2, x_3)^t$, $\Psi(x) = x_3(x_2 - x_1^2)^{-3/2}$. The gradient of the function $\Psi(\cdot)$ evaluated at x is

$$(\nabla \Psi)_x = \left(\frac{3x_1x_3}{(x_2 - x_1^2)^{5/2}}, -\frac{3}{2} \frac{x_3}{(x_2 - x_1^2)^{5/2}}, \frac{1}{(x_2 - x_1^2)^{3/2}} \right)^t.$$

Hence, using equation (42), the vector $D_{r_{n,E}}(\nabla \Psi)_{\theta \hat{r}_{n,E} + (1-\theta)r_{n,E}}$ is given by

$$\begin{aligned} &\left(\frac{3[\text{RCTM}_1(\alpha_n|x)]^2 \text{RCTM}_3(\alpha_n|x)}{[\widetilde{\text{RCTV}}(\alpha_n|x)]^{5/2}} (1 + O_{\mathbb{P}}(\sigma_n)), -\frac{3 \text{RCTM}_2(\alpha_n|x) \text{RCTM}_3(\alpha_n|x)}{[\widetilde{\text{RCTV}}(\alpha_n|x)]^{5/2}} (1 + O_{\mathbb{P}}(\sigma_n)), \right. \\ &\left. \frac{\text{RCTM}_3(\alpha_n|x)}{[\widetilde{\text{RCTV}}(\alpha_n|x)]^{3/2}} (1 + O_{\mathbb{P}}(\sigma_n)) \right)^t, \end{aligned}$$

where

$$\begin{aligned} \widetilde{\text{RCTV}}(\alpha_n|x) &= \text{RCTM}_2(\alpha_n|x)(1 + O_{\mathbb{P}}(\sigma_n)) - [\text{RCTM}_1(\alpha_n|x)]^2(1 + O_{\mathbb{P}}(\sigma_n))^2 \\ &= \text{RCTV}(\alpha_n|x) + O_{\mathbb{P}}(\sigma_n \text{RCTM}_2(\alpha_n|x)) + O_{\mathbb{P}}(\sigma_n [\text{RCTM}_1(\alpha_n|x)]^2). \end{aligned}$$

Since by assumption $\sigma_n [\text{RCTS}(\alpha_n|x)]^{2/3} \rightarrow 0$ and, by Proposition 1(ii), both $\text{RCTM}_2(\alpha_n|x)$ and $[\text{RCTM}_1(\alpha_n|x)]^2$ are $O([\text{RVaR}(\alpha_n|x)]^2)$, it appears that $\widetilde{\text{RCTV}}(\alpha_n|x) = \text{RCTV}(\alpha_n|x)(1 + o_{\mathbb{P}}(1))$. Using again Proposition 1(ii), it is then easy to check that $v_n D_{r_{n,E}}(\nabla \Psi)_{\theta \hat{r}_{n,E} + (1-\theta)r_{n,E}} \xrightarrow{P} v$ with $v_n = [\text{RCTS}(\alpha_n|x)]^{-5/3}$ and

$$v = \left(\frac{3(1 - 3\gamma_+(x))^{2/3}}{(1 - \gamma_+(x))^2}, -\frac{3}{2} \frac{(1 - 3\gamma_+(x))^{2/3}}{1 - 2\gamma_+(x)}, \frac{\gamma_+^2(x)(1 - 3\gamma_+(x))^{2/3}}{(1 - \gamma_+(x))^2(1 - 2\gamma_+(x))} \right)^t.$$

Applying Lemma 7 and remarking that $v^t \Sigma_E^{(1)}(x)v = \vartheta_{\text{RCTS},1}(\gamma_+(x))$, $v^t \Sigma_E^{(2)}(x)v = \vartheta_{\text{RCTS},2}(\gamma_+(x))$ concludes the proof. \blacksquare

Proof of Corollary 3. (i) First remark that $\hat{y}_{b,n}^*(x) = \Psi(\widehat{\text{RCTM}}_{b,n}(\alpha_n|x))$, with $\Psi(x) = x^{1/b}$ for $x \geq 0$. Since $(\nabla \Psi)_x = b^{-1}x^{1/b-1}$, one has for $\theta \in (0, 1)$ and in the situation $y^*(x) < \infty$ that

$$\begin{aligned} &\text{RCTM}_b(\alpha_n|x)(\nabla \Psi)_{\theta \widehat{\text{RCTM}}_{b,n}(\alpha_n|x) + (1-\theta)\text{RCTM}_b(\alpha_n|x)} \\ &= \frac{\text{RVaR}(\alpha_n|x)}{b} [\theta \widehat{\text{RCTM}}_{b,n}(\alpha_n|x) + (1-\theta)\text{RCTM}_b(\alpha_n|x)]^{1/b-1} = \frac{\text{RVaR}(\alpha_n|x)}{b} (1 + o_{\mathbb{P}}(1)), \end{aligned}$$

from Proposition 1(ii) and since, for all $\theta \in (0, 1)$,

$$\theta \widehat{\text{RCTM}}_{b,n}(\alpha_n|x) + (1-\theta)\text{RCTM}_b(\alpha_n|x) = \text{RCTM}_b(\alpha_n|x) (1 + O_{\mathbb{P}}(\sigma_n)),$$

from Theorem 2 and where $\sigma_n := (n \underline{\ell}^p \alpha_n)^{-1/2} \rightarrow 0$. Hence, the assumptions of Lemma 7 are satisfied with $v_n = 1/\text{RVaR}(\alpha_n|x)$ and $v = 1/b$ and thus,

$$(n \underline{\ell}^p \alpha_n)^{-1/2} [\text{RVaR}(\alpha_n|x)]^{-1} \left(\hat{y}_{b,n}^*(x) - [\text{RCTM}_b(\alpha_n|x)]^{1/b} \right) \xrightarrow{d} \mathcal{N}(0, \|\mathcal{K}\|^2 / (b^2 g(x))),$$

since $v^t \Sigma_{\{b\}}^{(1)}(x) v = v^t \Sigma_{\{b\}}^{(2)}(x) v = b^{-2}$. We conclude the proof by remarking that, from Lemma 3,

$$(n \underline{\ell}^p \alpha_n)^{1/2} [\text{RVaR}(\alpha_n|x)]^{-1} \left([\text{RCTM}_b(\alpha_n|x)]^{1/b} - \text{RVaR}(\alpha_n|x) \right) = O \left((n \underline{\ell}^p \alpha_n)^{1/2} \Delta_a(\text{RVaR}(\alpha_n|x)|x) \right)$$

and that $\text{RVaR}(\alpha_n|x) \rightarrow y^*(x)$ as $n \rightarrow \infty$.

(ii) Let us now consider the bias term in the situation $\gamma(x) < 0$:

$$\begin{aligned} (n \underline{\ell}^p \alpha_n)^{1/2} (\text{RVaR}(\alpha_n|x) - y^*(x)) &= (n \underline{\ell}^p \alpha_n)^{1/2} \Delta_a(\text{RVaR}(\alpha_n|x)|x) \text{RVaR}(\alpha_n|x) \frac{\text{RVaR}(\alpha_n|x) - y^*(x)}{a(\text{RVaR}(\alpha_n|x)|x)} \\ &= \frac{y^*(x)}{\gamma(x)} (n \underline{\ell}^p \alpha_n)^{1/2} \Delta_a(\text{RVaR}(\alpha_n|x)|x) (1 + o(1)), \end{aligned}$$

since $\text{RVaR}(\alpha_n|x) \rightarrow y^*(x)$ and $(\text{RVaR}(\alpha_n|x) - y^*(x))/a(\text{RVaR}(\alpha_n|x)|x) \rightarrow 1/\gamma(x)$ as $n \rightarrow \infty$ in view of (1.2.14) in [28, Lemma 1.2.9]. It is thus clear that the bias term tends to 0 under the assumption $(n \underline{\ell}^p \alpha_n)^{1/2} \Delta_a(\text{RVaR}(\alpha_n|x)|x) \rightarrow 0$ as $n \rightarrow \infty$. ■

$\vartheta_{\text{RCTE},1}(z)$	$\frac{(1-z)^2}{1-2z}$
$\vartheta_{\text{RCTE},2}(z)$	$(1-z)^2$
$\vartheta_{\text{RCTV},1}(z)$	$\frac{1-7z+9z^2+15z^3-6z^4}{(1-z)^2(1-2z)(1-3z)(1-4z)}$
$\vartheta_{\text{RCTV},2}(z)$	$\frac{(1+z)^2}{(1-z)^2}$
$\vartheta_{\text{RCTS},1}(z)$	$\frac{(1-3z)^{1/3}(9-198z+1749z^2-7890z^3+18979z^4-22746z^5+10415z^6+810z^7-1296z^8-216z^9)}{4(1-z)^4(1-2z)^2(1-4z)(1-5z)(1-6z)}$
$\vartheta_{\text{RCTS},2}(z)$	$\frac{(1-3z)^{4/3}(3-6z-z^2)^2}{4(1-z)^4(1-2z)^2}$
$\tilde{\vartheta}_{\text{RCTE},1}(z)$	$\frac{(1-z)^2}{1-2z}$
$\tilde{\vartheta}_{\text{RCTE},2}(z)$	$(1-z)^2$
$\tilde{\vartheta}_{\text{RCTV},1}(z)$	$\frac{(1-z)^2(1-2z)(1-7z+9z^2+15z^3-6z^4)}{z^4(1-3z)(1-4z)}$
$\tilde{\vartheta}_{\text{RCTV},2}(z)$	$\frac{(1-z)^2(1-2z)^2(1+z)^2}{z^4}$
$\tilde{\vartheta}_{\text{RCTS},1}(z)$	$\frac{(1-3z)^{1/3}(9-198z+1749z^2-7890z^3+18979z^4-22746z^5+10415z^6+810z^7-1296z^8-216z^9)}{4z^4(1-4z)(1-5z)(1-6z)}$
$\tilde{\vartheta}_{\text{RCTS},2}(z)$	$\frac{(1-3z)^{4/3}(3-6z-z^2)^2}{4z^4}$

Table 2: Asymptotic variances.

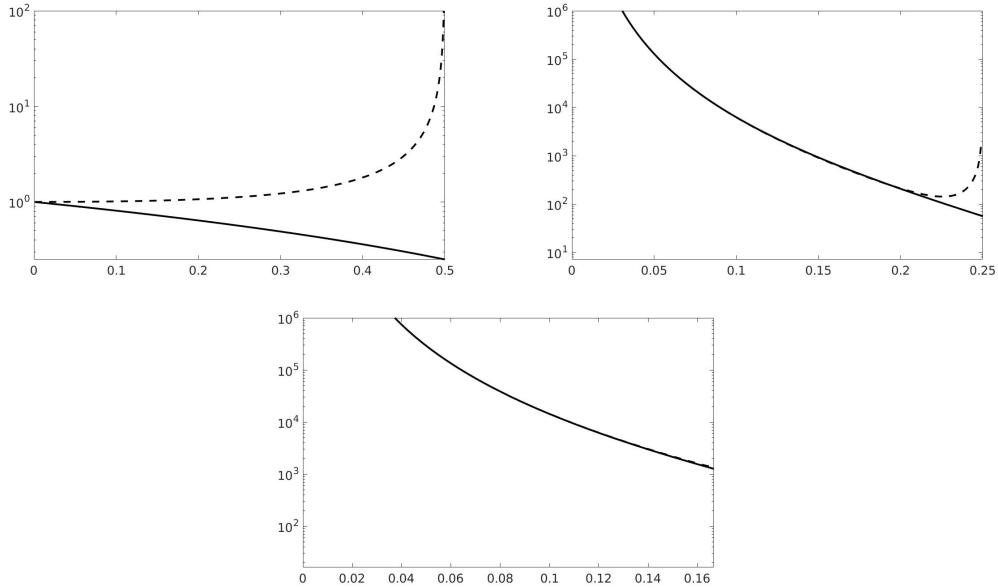


Figure 1: Asymptotic variances $z \mapsto \tilde{\vartheta}_{\text{RCTE},\bullet}(z)$ (top left), $z \mapsto \tilde{\vartheta}_{\text{RCTV},\bullet}(z)$ (top right) and $z \mapsto \tilde{\vartheta}_{\text{RCTS},\bullet}(z)$ (bottom) in a logarithmic scale. The case $k/h \rightarrow 0$ is depicted with a solid line while the case $h/k \rightarrow 0$ is depicted with a dashed line. In case of the RCTS, the two curves are almost superimposed.

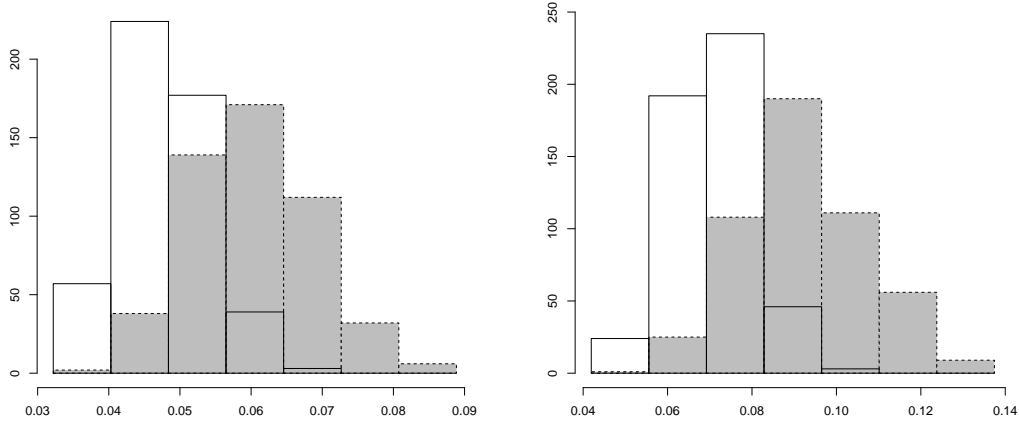


Figure 2: Comparison between the L^1 - error distributions associated with $\hat{y}_{7,n}^*$ computed on $N = 500$ samples of size $n = 500$. White bars: oracle strategy, grey bars: data-driven strategy, left: shape parameter $\theta_1(\cdot)$, right: shape parameter $\theta_2(\cdot)$.

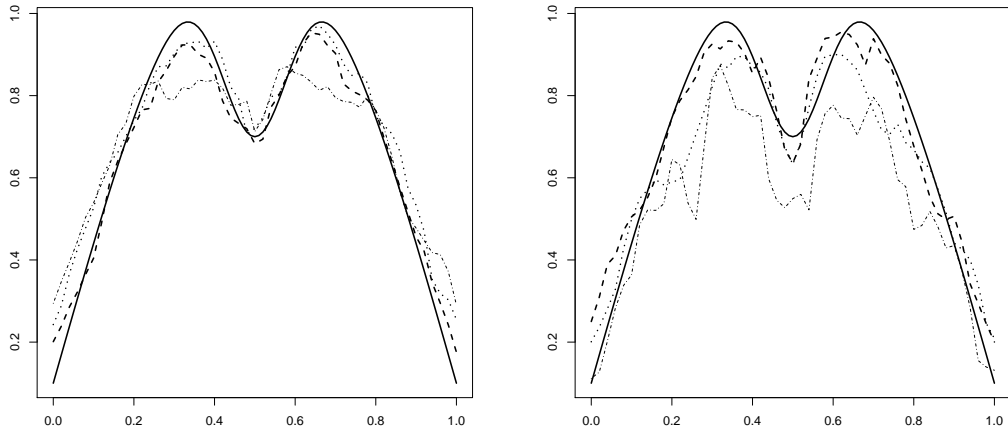


Figure 3: The frontier (solid line) and its estimations by $\hat{y}_{7,n}^*$ computed on the $N = 500$ samples of size $n = 500$ (left: shape parameter $\theta_1(\cdot)$, right: shape parameter $\theta_2(\cdot)$). For the mean L^1 - errors: best estimation (dotted line), worst estimation (dashed-dotted line) and median estimation (dashed line).

	$\theta(\cdot) = \theta_1(\cdot)$	$\theta(\cdot) = \theta_2(\cdot)$
$\widehat{\text{RVaR}}_n$	0.081 [0.045, 0.131]	0.131 [0.074, 0.199]
$\hat{y}_{1,n}^*$	0.108 [0.069, 0.169]	0.138 [0.076, 0.232]
$\hat{y}_{2,n}^*$	0.073 [0.047, 0.112]	0.107 [0.071, 0.177]
$\hat{y}_{3,n}^*$	0.065 [0.041, 0.098]	0.100 [0.062, 0.159]
$\hat{y}_{4,n}^*$	0.062 [0.039, 0.091]	0.096 [0.058, 0.149]
$\hat{y}_{5,n}^*$	0.061 [0.037, 0.089]	0.094 [0.056, 0.144]
$\hat{y}_{6,n}^*$	0.060 [0.037, 0.089]	0.093 [0.055, 0.140]
$\hat{y}_{7,n}^*$	0.059 [0.037, 0.088]	0.092 [0.054, 0.137]
$\hat{y}_{8,n}^*$	0.077 [0.032, 0.143]	0.104 [0.055, 0.171]
$\hat{y}_{9,n}^*$	0.070 [0.032, 0.131]	0.100 [0.053, 0.164]
$\hat{y}_{10,n}^*$	0.066 [0.033, 0.121]	0.098 [0.052, 0.158]
$\hat{y}_n^{(*,mc)}$	0.082 [0.054, 0.137]	0.141 [0.091, 0.202]
$\hat{y}_n^{(*,mv)}$	0.083 [0.051, 0.143]	0.140 [0.089, 0.203]
$\hat{y}_n^{(*,gj)}$	0.091 [0.043, 0.161]	0.246 [0.175, 0.326]

Table 3: Mean L^1 – errors and [minimum, maximum] L^1 – errors associated with the frontier estimators. Estimators $\widehat{\text{RVaR}}_n, \hat{y}_{1,n}^*, \dots, \hat{y}_{10,n}^*$ are computed with the hyper-parameters selected by the data-driven strategy. Results obtained with estimators $\hat{y}_n^{(*,mc)}, \hat{y}_n^{(*,mv)}$ and $\hat{y}_n^{(*,gj)}$ are reported from [24]. The best results are emphasized.

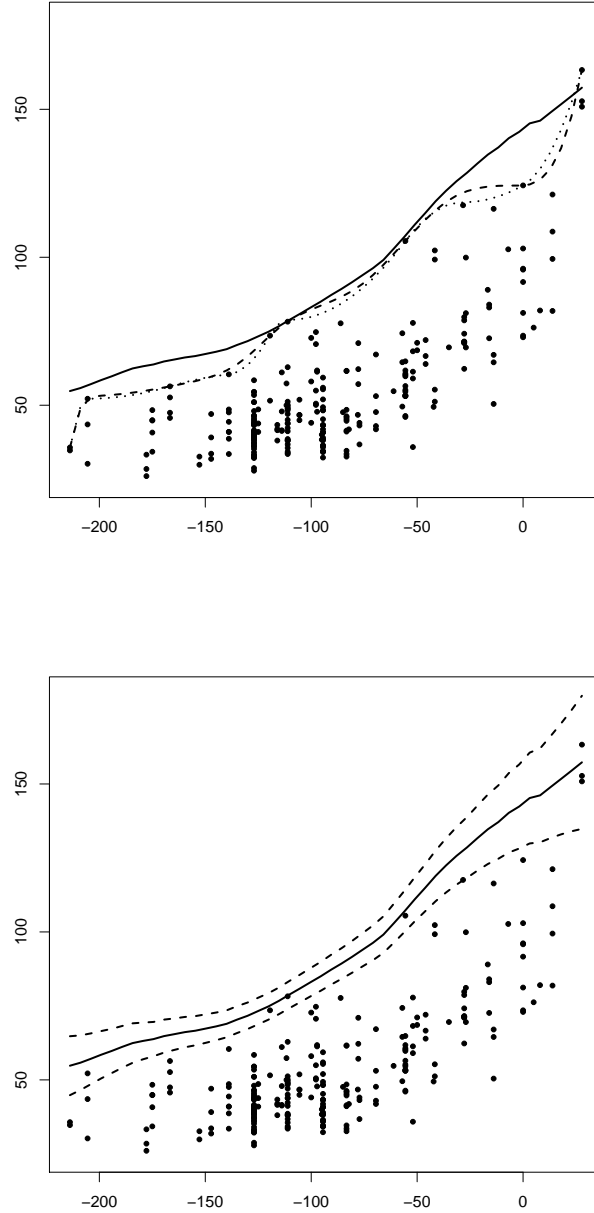


Figure 4: Scatterplot of the 254 nuclear reactor's data together with the frontier estimators. Top: $\hat{y}_{7,n}^*$ (solid line), CS-B (dashed line) and QS-B (dotted line). Bottom: $\hat{y}_{7,n}^*$ (solid line) and the associated pointwise asymptotic confidence intervals (dashed lines).

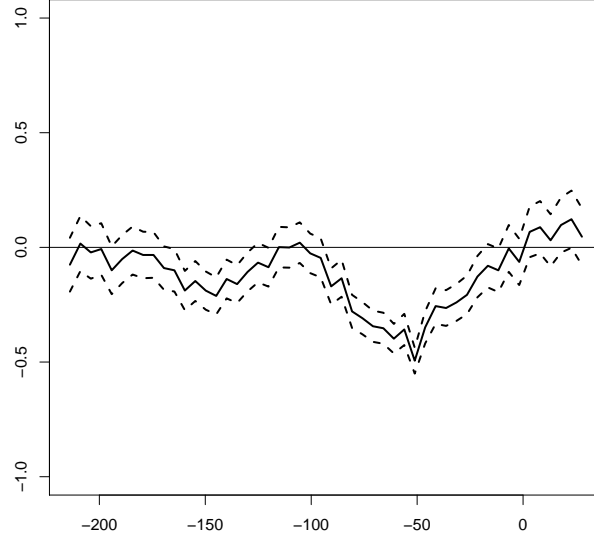


Figure 5: $\hat{\gamma}_n^{RP,1}$ (solid line) and the associated pointwise asymptotic confidence intervals (dashed lines) on the nuclear reactor's data.

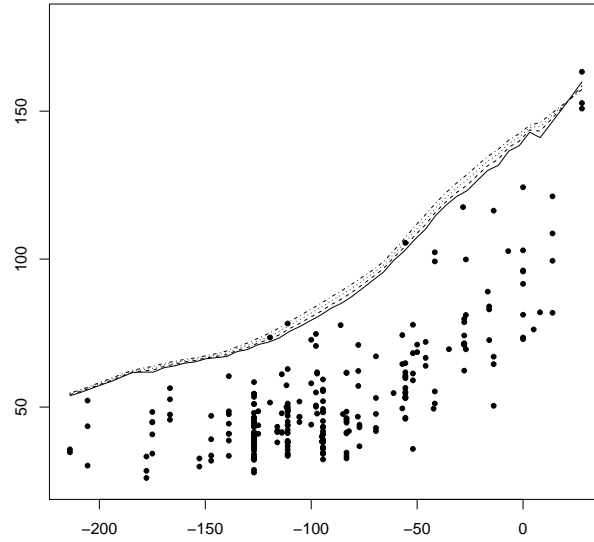


Figure 6: Scatterplot of the 254 nuclear reactor's data together with the frontier estimators $\hat{y}_{4,n}^*$, $\hat{y}_{5,n}^*$, $\hat{y}_{6,n}^*$ and $\hat{y}_{7,n}^*$. The four estimators are nearly superimposed.